```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          November 11, 2005, 09:31:07 ; Search time 3864 Seconds
                                         (without alignments)
                                         3856.105 Million cell updates/sec

Title:           US-09-522-753-5
Perfect score:   13215
Sequence:        1 MSGSTQLVAQTWRATEPRYP..........WDEEPKPLLCSQYETLSDSE 2517

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:      8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09522753/runat_02112005_173640_20785/app_query.fasta_1.2695
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09522753_@CGN_1_1_1431_@runat_02112005_173640_20785 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_16Dec04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002as:*
                 7:  geneseqn2002bs:*
                 8:  geneseqn2003as:*
                 9:  geneseqn2003bs:*
                 10:  geneseqn2003cs:*
                 11:  geneseqn2003ds:*
                 12:  geneseqn2004as:*
                 13:  geneseqn2004bs:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result           Query
   No.   Score   Match Length DB  ID                       Description
 ------------------------------------------------------------------------
     1   13215   100.0   7554  12  ADJ92815                 Adj92815 Human co-
     2   13215   100.0   8561  12  ADG86298                 Adg86298 Human SMR
     3   13215   100.0   8561  12  ADN04304                 Adn04304 Antipsori
     4 13178.5    99.7   8564  3  AAC74783                 Aac74783 Human ORF
     5   13145    99.5   8667  11  ACN44283                 Acn44283 Human mRN
     6 13119.5    99.3   8533  13  ADQ84524                 Adq84524 Human tum
     7 13119.5    99.3   8533  13  ACN39603                 Acn39603 Tumour-as
     8   12978    98.2   7521  8  ACA62250                 Aca62250 Human nuc
     9   12978    98.2   7524  10  ADL13812                 Adl13812 Osteoarth
    10   12978    98.2   8686  8  ACA62249                 Aca62249 cDNA enco
    11   12978    98.2   8686  10  ADL13811                 Adl13811 Osteoarth
    12   12978    98.2   8686  12  ADG86290                 Adg86290 Human SMR
    13   12978    98.2   8686  12  ADQ18920                 Adq18920 Human sof
    14   12922    97.8   9053  12  ADL12577                 Adl12577 Human ste
c   15 12845.5    97.2   9079  12  ADQ23294                 Adq23294 Human sof
    16 10832.5    82.0   7386  8  ACA62452                 Aca62452 Mouse nuc
    17 10832.5    82.0   8544  8  ACA62451                 Aca62451 cDNA enco
    18  9691.5    73.3   7534  11  ACN44281                 Acn44281 Mouse mRN
    19    7885    59.7   5989  6  ABK84305                 Abk84305 Human cDN
    20    4612    34.9   2930  10  ADC35130                 Adc35130 Human bre
    21    4612    34.9   2930  12  ADG86301                 Adg86301 Human SMR
    22    4187    31.7   7780  3  AAA60630                 Aaa60630 HNRCR nuc
    23  4165.5    31.5   7914  12  ADQ83959                 Adq83959 Human tum
    24  4165.5    31.5   7914  13  ADQ87683                 Adq87683 Human tum
    25  4147.5    31.4   7940  8  ABZ34833                 Abz34833 Coding se
    26  4147.5    31.4   7940  13  ADQ89791                 Adq89791 Antagonis
    27    4043    30.6   7900  3  AAA60629                 Aaa60629 Human HNR
    28    3649    27.6  91141  11  ACN44280                 Acn44280 Mouse gen
    29    2951    22.3 233380  11  ACN44282                 Acn44282 Human gen
    30    2184    16.5   2745  12  ADJ92816                 Adj92816 Human co-
    31  2133.5    16.1 220756  12  ADG86300                 Adg86300 Human SMR
c   32  1871.5    14.2   2336  10  ADE31306                 Ade31306 Human dia
    33    1391    10.5   2914  13  ADR07585                 Adr07585 Full leng
    34    1113     8.4    650  8  ACA57401                 Aca57401 Human adi
    35    1099     8.3  10410  13  ADS96633                 Ads96633 Drosophil
    36    1083     8.2  10910  4  ABL03131                 Abl03131 Drosophil
    37    1076     8.1    872  3  AAA02670                 Aaa02670 Human col
    38  1073.5     8.1  10900  13  ADQ89593                 Adq89593 Antagonis
    39   977.5     7.4   2519  12  ADP48942                 Adp48942 Human-yea
    40     960     7.3    555  8  ACA57524                 Aca57524 Human adi
c   41     880     6.7    718  10  ADE76358                 Ade76358 Human BSK
c   42     845     6.4    710  4  AAI97539                 Aai97539 Human neu
    43   774.5     5.9    673  8  ACA57523                 Aca57523 Human adi
    44   746.5     5.6    520  6  ABK44934                 Abk44934 cDNA enco
c   45   726.5     5.5  33529  5  AAS17367                 Aas17367 DNA seque
```

```
                                ALIGNMENTS

RESULT 1
ADJ92815
ID   ADJ92815 standard; DNA; 7554 BP.
XX
AC   ADJ92815;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human co-repressor SMRT DNA.
XX
KW   Nuclear receptor; therapy; diabetes; hyperlipidaemia; obesity;
KW   atherosclerosis; human; co-repressor; gene; ds.
XX
OS   Homo sapiens.
XX
PN   US2003228607-A1.
XX
PD   11-DEC-2003.
XX
PF   14-APR-2003; 2003US-00414692.
XX
PR   15-APR-2002; 2002US-0372650P.
XX
PA   (WAGN/) WAGNER B L.
PA   (SCHU/) SCHULMAN I G.
XX
PI   Wagner BL,  Schulman IG;
XX
DR   WPI; 2004-167207/16.
```

```
XX
PT   Identifying compounds that bind to nuclear receptor and exhibit cell type
PT   specific actions, and useful for treating hyperlipidemia, obesity and
PT   diabetes.
XX
PS   Claim 5; SEQ ID NO 14; 99pp; English.
XX
CC   The invention relates to screening methods for identifying compounds that
CC   bind to nuclear receptor and exhibit cell type specific actions. The
CC   invention relates to modulators having an improved therapeutic profile.
CC   The method is useful for identifying compounds that bind to a nuclear
CC   receptor and exhibit cell type specific actions. It is also useful for
CC   identifying modulators of nuclear receptors that are useful in treating
CC   diseases e.g., diabetes, hyperlipidaemia, obesity, atherosclerosis, etc.
CC   The present sequence is human co-repressor DNA used to illustrate the
CC   method of the invention.
XX
SQ   Sequence 7554 BP; 1646 A; 2732 C; 2228 G; 948 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0              Length:        7554
Score:                    13215.00        Matches:       2517
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              100.00%        Indels:        0
DB:                       12             Gaps:          0

US-09-522-753-5 (1-2517) x ADJ92815 (1-7554)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 900

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 960

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1020

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1080

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1140

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1200

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1260

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1320

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1380

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1440

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1500

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1560

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1620

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1680

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1740

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1800

Qy         601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1860

Qy         621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1920

Qy         641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1980

Qy         661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2040

Qy         681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2100

Qy         701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2160

Qy         721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2220

Qy         741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2280

Qy         761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2340

Qy         781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2400

Qy         801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2460

Qy         821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2520

Qy         841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GAGGAGCAGAAGCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2580

Qy         861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2640

Qy         881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2700

Qy         901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCAGGACAGCGACTCCAGTGCTACC 2760

Qy         921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2820

Qy         941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        2821 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2880

Qy         961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2940

Qy         981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3000

Qy        1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3060

Qy        1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 GGCAAGAGCAGGAGCCCGGCACCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3120

Qy        1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3180

Qy        1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3240

Qy        1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3300

Qy        1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3360

Qy        1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3420

Qy        1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3421 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3480

Qy        1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3481 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3540

Qy        1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3541 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3600

Qy        1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3660

Qy        1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3661 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3720

Qy        1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3780

Qy        1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3840

Qy        1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3900

Qy        1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3960
```

```
Qy       1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4020

Qy       1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4080

Qy       1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4140

Qy       1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4200

Qy       1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4260

Qy       1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4320

Qy       1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4380

Qy       1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4440

Qy       1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4500

Qy       1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4560

Qy       1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4620

Qy       1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4680

Qy       1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4740

Qy       1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4800

Qy       1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4860

Qy       1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4920

Qy       1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4980

Qy       1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5040

Qy       1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5100

Qy       1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5101 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5160

Qy       1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5161 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5220

Qy       1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5221 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5280

Qy       1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5281 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5340

Qy       1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5341 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5400

Qy       1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5401 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5460

Qy       1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5461 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5520

Qy       1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5521 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5580

Qy       1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5581 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5640

Qy       1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5641 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5700

Qy       1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5701 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5760

Qy       1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5761 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5820

Qy       1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5821 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5880

Qy       1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5881 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5940

Qy       1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5941 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6000

Qy       2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6001 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6060

Qy       2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6061 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6120

Qy       2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6121 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6180

Qy        2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6181 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6240

Qy        2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6241 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6300

Qy        2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6301 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6360

Qy        2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6361 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6420

Qy        2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6421 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6480

Qy        2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6481 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6540

Qy        2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6541 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6600

Qy        2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6601 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6660

Qy        2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6661 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6720

Qy        2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6721 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6780

Qy        2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6781 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6840

Qy        2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6841 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6900

Qy        2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6901 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6960

Qy        2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6961 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7020

Qy        2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7021 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7080

Qy        2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7081 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7140

Qy        2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7141 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7200

Qy        2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7201 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7260

Qy        2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7261 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7320

Qy        2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7321 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7380

Qy        2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7381 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7440

Qy        2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7441 CCACCGGGCCTCCCCGCGGGCAGCGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7500

Qy        2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7501 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7551

RESULT 2
ADG86298
ID   ADG86298 standard; cDNA; 8561 BP.
XX
AC   ADG86298;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human SMRT encoding cDNA SEQ ID NO:12.
XX
KW   SMRT; silencing mediator for retinoid and thyroid hormone action;
KW   SMRT inhibitor; cytostatic; antiinflammatory; antiarthritic;
KW   antirheumatic; antisense therapy; inflammatory disorder;
KW   rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;
KW   breast cancer; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2. .7555
FT                   /*tag=  a
FT                   /product= "SMRT"
XX
PN   WO2003106645-A2.
XX
PD   24-DEC-2003.
XX
PF   17-JUN-2003; 2003WO-US018923.
XX
PR   17-JUN-2002; 2002US-00174014.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Bennett CF,  Freier SM,  Dobie KW;
XX
DR   WPI; 2004-082184/08.
DR   P-PSDB; ADG86299.
DR   GENBANK; NM_006312.
XX
PT   Novel antisense compound targeted to nucleic acid encoding SMRT
PT   (silencing mediator for retinoid and thyroid hormone action), useful for
PT   treating animal having disease associated with SMRT such as cancer,
PT   rheumatoid arthritis.
XX
PS   Example 15; SEQ ID NO 12; 260pp; English.
XX
CC   The present invention describes a compound (I) 8-50 nucleobases in length
CC   targeted to a nucleic acid molecule encoding SMRT (silencing mediator for
CC   retinoid and thyroid hormone action), where (I) specifically hybridises
CC   with the nucleic acid molecule encoding SMRT and inhibits expression of
CC   SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a
CC   preferred target region on nucleic acid molecule encoding SMRT. Also
CC   described is a composition (II) comprising (I) and a carrier or diluent.
CC   (I) and (II) have cytostatic, antiinflammatory, antiarthritic and
CC   antirheumatic activities, and can be used in antisense therapy, and as
CC   SMRT expression inhibitors. (I) is useful for inhibiting the expression
CC   of SMRT in cells or tissues. (I) is also useful for treating an animal
CC   having a disease or condition associated with SMRT, e.g., inflammatory
CC   disorder such as rheumatoid arthritis; or a hyperproliferative disorder
CC   such as cancer chosen from leukaemia and breast cancer, by inhibiting the
CC   expression of SMRT. (I) is useful for diagnostics, therapeutics,
CC   prophylaxis and as research reagents and kits. The present sequence
CC   encodes human SMRT, which is used in an example from the present
CC   invention. N.B. The present sequence is designated as SEQ ID NO:11 in
CC   example 15 but corresponds to SEQ ID NO:12 in the Sequence Listing.
XX
SQ   Sequence 8561 BP; 1862 A; 3033 C; 2525 G; 1141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0            Length:          8561
Score:                  13215.00      Matches:         2517
Percent Similarity:     100.00%      Conservative:    0
Best Local Similarity:  100.00%      Mismatches:      0
Query Match:            100.00%      Indels:          0
DB:                     12           Gaps:            0

US-09-522-753-5 (1-2517) x ADG86298 (1-8561)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 181

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 901

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 961

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1021

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1501

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801
```

```
Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy        781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2401

Qy        801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2461

Qy        821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2521

Qy        841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2522 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2581

Qy        861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2641

Qy        881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2701

Qy        901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2761

Qy        921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2821

Qy        941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2881

Qy        961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2941

Qy        981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2942 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3001

Qy       1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3061

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3121

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3181

Qy       1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3241

Qy       1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3301

Qy       1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3361

Qy       1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3421

Qy       1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3481

Qy       1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3541

Qy       1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3601

Qy       1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3661

Qy       1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3721

Qy       1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3781

Qy       1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3782 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3841

Qy       1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3842 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3901

Qy       1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3902 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3961

Qy       1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3962 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4021

Qy          1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4022 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4081

Qy          1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4082 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4141

Qy          1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4142 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4201

Qy          1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4261

Qy          1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4262 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4321

Qy          1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4322 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4381

Qy          1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4382 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4441

Qy          1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4442 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4501

Qy          1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4502 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4561

Qy          1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4562 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4621

Qy          1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4622 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4681

Qy          1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4682 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4741

Qy          1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4742 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4801

Qy          1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4802 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4861

Qy          1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4862 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4921

Qy          1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4922 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4981

Qy          1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4982 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5041

Qy          1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          5042 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5101

Qy          1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5102 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5161

Qy          1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5162 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5221

Qy          1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5222 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5281

Qy          1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5282 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5341

Qy          1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5342 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5401

Qy          1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5402 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5461

Qy          1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5462 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5521

Qy          1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5522 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5581

Qy          1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5582 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5641

Qy          1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5642 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5701

Qy          1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5702 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5761

Qy          1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5762 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5821

Qy          1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5822 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5881

Qy          1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5882 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5941

Qy          1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5942 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6001

Qy          2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6002 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6061

Qy          2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6062 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6121

Qy          2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6122 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6181
```

```
Qy       2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6182 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6241

Qy       2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6242 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6301

Qy       2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6302 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6361

Qy       2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6362 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6421

Qy       2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6422 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6481

Qy       2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6482 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6541

Qy       2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6542 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6601

Qy       2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6602 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6661

Qy       2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6662 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6721

Qy       2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6722 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6781

Qy       2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6782 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6841

Qy       2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6842 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6901

Qy       2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6902 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6961

Qy       2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6962 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7021

Qy       2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7022 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7081

Qy       2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7082 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7141

Qy       2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7142 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7201

Qy       2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7202 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7261

Qy       2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7262 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7321

Qy       2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7322 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7381

Qy       2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7382 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7441

Qy       2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7442 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7501

Qy       2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7502 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7552

RESULT 3
ADN04304
ID   ADN04304 standard; cDNA; 8561 BP.
XX
AC   ADN04304;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic cDNA sequence #352.
XX
KW   ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   P-PSDB; ADN04305.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 1; SEQ ID NO 698; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polynucleotides of the invention.
XX
SQ   Sequence 8561 BP; 1862 A; 3033 C; 2525 G; 1141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0             Length:        8561
Score:                  13215.00       Matches:       2517
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     12            Gaps:          0

US-09-522-753-5 (1-2517) x ADN04304 (1-8561)
```

```
Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 181

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 901

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 961

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1021

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1501

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy     741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy     761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy     781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2342 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2401

Qy     801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2402 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2461

Qy     821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2462 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2521

Qy     841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2522 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2581

Qy     861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2582 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2641

Qy     881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2642 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2701

Qy     901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2702 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2761

Qy     921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2762 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2821

Qy     941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2822 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2881

Qy     961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2941

Qy     981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2942 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3001

Qy    1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3002 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3061

Qy    1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3062 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3121

Qy    1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3122 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3181

Qy    1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3182 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3241

Qy    1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3242 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3301

Qy    1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3302 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3361

Qy    1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3362 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3421

Qy    1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3422 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3481

Qy    1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3482 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3541

Qy    1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3542 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3601

Qy    1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3602 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3661

Qy    1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3662 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3721

Qy    1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3722 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3781

Qy    1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3782 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3841

Qy    1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3842 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3901

Qy    1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3902 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3961

Qy    1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3962 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4021

Qy    1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4022 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4081

Qy    1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4082 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4141

Qy    1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4142 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4201

Qy    1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4261

Qy    1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4262 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4321

Qy    1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4322 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4381
```

```
Qy         1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4382 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4441

Qy         1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4442 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4501

Qy         1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4502 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4561

Qy         1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4562 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4621

Qy         1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4622 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4681

Qy         1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4682 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4741

Qy         1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4742 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4801

Qy         1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4802 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4861

Qy         1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4862 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4921

Qy         1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4922 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4981

Qy         1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4982 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5041

Qy         1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5042 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5101

Qy         1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5102 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5161

Qy         1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5162 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5221

Qy         1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5222 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5281

Qy         1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5282 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5341

Qy         1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5342 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5401

Qy         1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5402 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5461

Qy         1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5462 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5521

Qy         1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5522 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5581

Qy         1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5582 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5641

Qy         1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5642 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5701

Qy         1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5702 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5761

Qy         1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5762 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5821

Qy         1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5822 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5881

Qy         1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5882 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5941

Qy         1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5942 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6001

Qy         2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6002 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6061

Qy         2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6062 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6121

Qy         2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6122 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6181

Qy         2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6182 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6241

Qy         2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6242 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6301

Qy         2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6302 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6361

Qy         2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6362 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6421

Qy         2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6422 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6481

Qy         2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6482 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6541

Qy         2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6542 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6601

Qy       2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6602 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6661

Qy       2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6662 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6721

Qy       2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6722 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6781

Qy       2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6782 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6841

Qy       2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6842 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6901

Qy       2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6902 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6961

Qy       2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6962 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7021

Qy       2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7022 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7081

Qy       2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7082 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7141

Qy       2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7142 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7201

Qy       2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7202 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7261

Qy       2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7262 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7321

Qy       2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7322 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7381

Qy       2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7382 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7441

Qy       2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7442 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7501

Qy       2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7502 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7552

RESULT 4
AAC74783
ID   AAC74783 standard; cDNA; 8564 BP.
XX
AC   AAC74783;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF338 polynucleotide sequence SEQ ID NO:675.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US008621.
XX
PR   31-MAR-1999;   99US-0127607P.
PR   02-APR-1999;   99US-0127636P.
PR   05-APR-1999;   99US-0127728P.
PR   30-MAR-2000; 2000US-00540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB40574.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease.
XX
PS   Claim 5; Page 779-784; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC   anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC   cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC   dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC   antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC   sequences can be used for determining the presence of or predisposition
CC   to, or preventing or treating pathological conditions associated with an
CC   ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC   proteins in gene therapy vectors. The proteins and nucleic acids may be
CC   used to treat cancers, proliferative disorders, neurodegenerative
CC   disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC   diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC   storage, systemic lupus erythematosus, severe combined immunodeficiency
CC   (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC   disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC   cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC   enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ   Sequence 8564 BP; 1859 A; 3034 C; 2532 G; 1139 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0            Length:          8564
Score:                  13178.50      Matches:         2512
Percent Similarity:     99.80%       Conservative:  1
Best Local Similarity:  99.76%       Mismatches:    4
```

```
Query Match:             99.72%         Indels:        1
DB:                      3              Gaps:          1

US-09-522-753-5 (1-2517) x AAC74783 (1-8564)

Qy             1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy            21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy            41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 180

Qy            61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy            81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy           101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360

Qy           121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy           141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy           161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy           181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy           201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy           221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy           241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy           261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy           281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                 |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db           841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 900

Qy           301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                 ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db           901 CGCTATGACCAGCTCATGGAGGCCTGGGAAAAAAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy           321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           961 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1020

Qy           341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                 |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db          1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGGGCAGGGTGGGCCAGCGGGGCAGTGGG 1080

Qy           361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1081 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1140

Qy           381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1141 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1200

Qy           401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1201 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1260

Qy           421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1261 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1320

Qy           441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1321 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1380

Qy           461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1381 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1440

Qy           481 SerLeuValArgArgSerTyrArgArgArgGlyLysSer---GlnGlnGlnGlnGlnGln 499
                 |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db          1441 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAGCAACAACAGCAG 1500

Qy           500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1501 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAG 1560

Qy           520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1561 AAAGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAG 1620

Qy           540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1621 AACGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGAC 1680

Qy           560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1681 GAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAA 1740

Qy           580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1741 GGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAG 1800

Qy           600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1801 CAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAA 1860

Qy           620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1861 ATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGG 1920

Qy           640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1921 ATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 1980

Qy           660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1981 CAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCG 2040

Qy           680 ArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProVal 699
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2041 CGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTG 2100

Qy           700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAG 2160

Qy         720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPro 739
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 GAGGCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCA 2220

Qy         740 AlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAla 759
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACGGAGGCCGCC 2280

Qy         760 LysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProPro 779
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 AAGGACACAGGGCAGAATGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGCCACCC 2340

Qy         780 ProGlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAla 799
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 CCAGGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCC 2400

Qy         800 SerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProPro 819
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 TCTGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCT 2460

Qy         820 ProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGlu 839
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 CCTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAG 2520

Qy         840 GlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGlu 859
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GGGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAG 2580

Qy         860 GluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAsp 879
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 GAGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGAC 2640

Qy         880 AlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGly 899
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 GCGGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGG 2700

Qy         900 SerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAla 919
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 AGCGGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCT 2760

Qy         920 ThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer 939
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 ACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCC 2820

Qy         940 ProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys 959
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 CCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAG 2880

Qy         960 ProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThr 979
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACC 2940

Qy         980 LysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaPro 999
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 AAAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCA 3000

Qy        1000 ProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerPro 1019
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCC 3060

Qy        1020 ArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAla 1039
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 CGGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCC 3120

Qy        1040 GlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPro 1059
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 CAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCC 3180

Qy        1060 ArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPro 1079
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 CGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCA 3240

Qy        1080 ProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPro 1099
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 CCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCA 3300

Qy        1100 ProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGlu 1119
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 CCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAG 3360

Qy        1120 ArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGlu 1139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 AGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAG 3420

Qy        1140 HisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLys 1159
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3421 CATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAA 3480

Qy        1160 LysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGly 1179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3481 AAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGG 3540

Qy        1180 ProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAla 1199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3541 CCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCT 3600

Qy        1200 LeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSer 1219
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 CTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCG 3660

Qy        1220 AspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyr 1239
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3661 GACAGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTAC 3720

Qy        1240 LysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArg 1259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 AAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGG 3780

Qy        1260 GluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSer 1279
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 GAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCC 3840

Qy        1280 TyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGly 1299
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 TATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGA 3900

Qy        1300 ProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGly 1319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 CCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGC 3960

Qy        1320 ArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArg 1339
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3961 AGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGA 4020

Qy        1340 HisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIle 1359
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4021 CACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATC 4080

Qy        1360 ProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLys 1379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4081 CCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAG 4140

Qy        1380 ArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThr 1399
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4141 CGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACG 4200

Qy        1400 GlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLys 1419
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4201 CAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAG 4260

Qy        1420 GluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeu 1439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4261 GAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTG 4320
```

```
Qy       1440 ProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyr 1459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4321 CCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTAC 4380

Qy       1460 AspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySer 1479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 GACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGC 4440

Qy       1480 ProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeu 1499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 CCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTG 4500

Qy       1500 GluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGly 1519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 GAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGG 4560

Qy       1520 GlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSer 1539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 GGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGC 4620

Qy       1540 ProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerPro 1559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 CCCCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCC 4680

Qy       1560 ValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAla 1579
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 GTGACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCA 4740

Qy       1580 SerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThr 1599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 TCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACC 4800

Qy       1600 ValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSer 1619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 GTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGT 4860

Qy       1620 GlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArg 1639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 GGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGC 4920

Qy       1640 GlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnPro 1659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 GGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCC 4980

Qy       1660 ThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeu 1679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 ACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTG 5040

Qy       1680 GluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThr 1699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5041 GAGAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACG 5100

Qy       1700 AlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSer 1719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5101 GCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCG 5160

Qy       1720 LeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHis 1739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5161 CTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACAC 5220

Qy       1740 LeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAla 1759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5221 CTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCC 5280

Qy       1760 TyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerPro 1779
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5281 TACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCA 5340

Qy       1780 GlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAsp 1799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5341 GGAGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGAC 5400

Qy       1800 ArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThr 1819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5401 CGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACG 5460

Qy       1820 ThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGly 1839
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5461 ACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGC 5520

Qy       1840 SerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGln 1859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5521 AGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAG 5580

Qy       1860 HisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHis 1879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5581 CACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCAC 5640

Qy       1880 AsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArg 1899
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db       5641 AACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGG 5700

Qy       1900 SerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCys 1919
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5701 TCCACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGC 5760

Qy       1920 ProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuPro 1939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5761 CCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCC 5820

Qy       1940 LysGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPhe 1959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5821 AAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTC 5880

Qy       1960 LeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySer 1979
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5881 CTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCG 5940

Qy       1980 GluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAla 1999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5941 GAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCG 6000

Qy       2000 LysAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSer 2019
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6001 AAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCG 6060

Qy       2020 AspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArg 2039
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6061 GACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGT 6120

Qy       2040 SerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProVal 2059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6121 TCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTG 6180

Qy       2060 SerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLys 2079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6181 AGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAG 6240

Qy       2080 SerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGlu 2099
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6241 AGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAG 6300

Qy       2100 AlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeu 2119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 GCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTG 6360

Qy       2120 LeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIle 2139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6361 CTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATC 6420

Qy       2140 SerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeu 2159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6421 AGTGAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTG 6480

Qy       2160 ProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgPro 2179
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6481 CCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCA 6540

Qy       2180 ProSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHis 2199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6541 CCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCAC 6600

Qy       2200 SerGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGlu 2219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6601 AGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAG 6660

Qy       2220 AspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSer 2239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6661 GACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGT 6720

Qy       2240 AlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySer 2259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6721 GCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCC 6780

Qy       2260 LysSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsn 2279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6781 AAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAAC 6840

Qy       2280 SerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArg 2299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6841 TCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGG 6900

Qy       2300 AsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIle 2319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6901 AATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATC 6960

Qy       2320 ThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsn 2339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6961 ACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAAC 7020

Qy       2340 MetGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu 2359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7021 ATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAG 7080

Qy       2360 SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAla 2379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7081 TCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCT 7140

Qy       2380 AlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGly 2399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7141 GCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGC 7200

Qy       2400 GlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro 2419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7201 GGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCG 7260

Qy       2420 GlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCys 2439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7261 GGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGC 7320

Qy       2440 AsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySer 2459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7321 AACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCC 7380

Qy       2460 ThrProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerPro 2479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7381 ACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCA 7440

Qy       2480 ProProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAsp 2499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7441 CCCCCACCGGGCCTCCCCGCGGGCAGCGGCCCCTCGCTGGCCCCCACCACGCCTGGGAC 7500

Qy       2500 GluGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7501 GAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7554

RESULT 5
ACN44283
```

```
ID   ACN44283 standard; cDNA; 8667 BP.
XX
AC   ACN44283;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human mRNA sequence hCT16428.
XX
KW   Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO2003073826-A2.
XX
PD   12-SEP-2003.
XX
PF   28-FEB-2003; 2003WO-US006235.
XX
PR   01-MAR-2002; 2002US-00087192.
XX
PA   (SAGR-) SAGRES DISCOVERY.
XX
PI   Morris DW;
XX
DR   WPI; 2003-328604/31.
XX
PT   Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT   comprises a nucleotide sequence.
XX
PS   Claim 1; SEQ ID NO 653; 0pp; English.
XX
CC   The present invention relates to novel DNA and protein sequences which
CC   are associated with carcinomas. The sequences are useful for: (i) for
CC   screening drug candidates; (ii) for screening of bioactive agent capable
CC   of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC   a bioactive agent capable of modulating the activity of CAP; (iv) for
CC   evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC   carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC   carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC   (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC   determining Carcinoma Associated (CA) gene copy number. In addition, the
CC   CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC   carcinoma including lymphoma. The present sequence is one such CA gene
CC   sequence. Note: This patent is an equivalent to basic patent
CC   US2002182586A1, for which no sequence data was published
XX
SQ   Sequence 8667 BP; 1860 A; 3066 C; 2575 G; 1166 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0              Length:          8667
Score:                  13145.00       Matches:         2509
Percent Similarity:     99.68%         Conservative:    1
Best Local Similarity:  99.64%         Mismatches:      7
Query Match:            99.47%         Indels:          2
DB:                     11             Gaps:            0

US-09-522-753-5 (1-2517) x ACN44283 (1-8667)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              |||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db        116 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 175

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 235

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 295

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 355
```

```
Qy      81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     356 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 415

Qy     101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     416 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 475

Qy     121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     476 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 535

Qy     141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     536 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 595

Qy     161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     596 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 655

Qy     181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     656 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 715

Qy     201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     716 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 775

Qy     221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     776 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 835

Qy     241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     836 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 895

Qy     261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     896 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 955

Qy     281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
           ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db     956 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1015

Qy     301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db    1016 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1075

Qy     321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1076 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1135

Qy     341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1136 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1195

Qy     361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1196 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1255

Qy     381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1256 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1315

Qy     401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1316 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1375

Qy     421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1376 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1435

Qy     441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1436 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1495

Qy     461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1496 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1555

Qy     481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1556 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1615

Qy     501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1616 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1675

Qy     521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1676 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1735

Qy     541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1736 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1795

Qy     561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1796 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1855

Qy     581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1856 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1915

Qy     601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1916 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1975

Qy     621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1976 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2035

Qy     641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2036 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2095

Qy     661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2096 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2155

Qy     681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2156 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2215

Qy     701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2216 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2275

Qy     721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2276 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2335

Qy     741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2336 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2395

Qy     761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2396 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2455

Qy     781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
           |||||||||||||||||| |||||||||||||||   |||   |||||||||||||||||
Db    2456 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCGGCCTC 2514

Qy     800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 TGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCC 2574

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2634

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2635 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2694

Qy        860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2695 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2754

Qy        880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2755 GGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2814

Qy        900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2815 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTAC 2874

Qy        920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2875 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2934

Qy        940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2935 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2994

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2995 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3054

Qy        980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3055 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3114

Qy       1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3115 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3174

Qy       1020 gGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGl 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3175 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCA 3234

Qy       1040 nLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProAr 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3235 GAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCG 3294

Qy       1060 gGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPr 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3295 TGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACC 3354

Qy       1080 oGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPr 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3355 TGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACC 3414

Qy       1100 oThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluAr 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3415 CACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAG 3474

Qy       1120 gGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHi 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3475 GCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCA 3534

Qy       1140 sAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLy 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3535 TGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAA 3594

Qy       1160 sLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPr 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3595 GCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCC 3654

Qy       1180 oProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLe 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3655 ACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCT 3714

Qy       1200 uGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAs 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3715 GGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGA 3774

Qy       1220 pSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLy 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3775 CAGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAA 3834

Qy       1240 sGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGl 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3835 GGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGA 3894

Qy       1260 uAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTy 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3895 GGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTA 3954

Qy       1280 rGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPr 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3955 TGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACC 4014

Qy       1300 oProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyAr 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4015 CCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAG 4074

Qy       1320 gAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHi 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4075 AGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACA 4134

Qy       1340 sSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePr 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4135 CAGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCC 4194

Qy       1360 oArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysAr 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4195 TCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCG 4254

Qy       1380 gGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGl 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4255 GGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCA 4314

Qy       1400 nAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGl 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4315 GGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGA 4374

Qy       1420 uAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPr 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4375 GGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCC 4434

Qy       1440 oLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAs 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4435 CCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGA 4494

Qy       1460 pThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPr 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4495 CACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCC 4554

Qy       1480 oGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGl 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4555 CGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGA 4614

Qy       1500 uArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGl 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4615 ACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGG 4674

Qy       1520 ySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPr 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4675 CTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCCGCGGCAGAGCCC 4734
```

```
Qy          1540 oLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVa 1560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4735 CCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGT 4794

Qy          1560 lThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSe 1580
                 ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4795 GACCACGCGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATC 4854

Qy          1580 rGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVa 1600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4855 CCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGT 4914

Qy          1600 lProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGl 1620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4915 GCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGG 4974

Qy          1620 yValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGl 1640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4975 CGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGG 5034

Qy          1640 yIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProTh 1660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5035 CATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCAC 5094

Qy          1660 rTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGl 1680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5095 CTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGA 5154

Qy          1680 uAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAl 1700
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db          5155 GAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGC 5214

Qy          1700 aThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLe 1720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5215 CACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCT 5274

Qy          1720 uAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLe 1740
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5275 GGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCT 5334

Qy          1740 uProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5335 GCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTA 5394

Qy          1760 rLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGl 1780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5395 CCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGG 5454

Qy          1780 yGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspAr 1800
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5455 AGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCG 5514

Qy          1800 gAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrTh 1820
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5515 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGAC 5574

Qy          1820 rValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySe 1840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5575 GGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAG 5634

Qy          1840 rSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHi 1860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5635 CAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCA 5694

Qy          1860 sSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAs 1880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5695 CTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAA 5754

Qy          1880 nThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSe 1900
                 ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db          5755 CACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGGTC 5814

Qy          1900 rThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPr 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5815 CACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCC 5874

Qy          1920 oLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLy 1940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5875 ACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAA 5934

Qy          1940 sGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLe 1960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5935 GGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCT 5994

Qy          1960 uAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGl 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5995 CGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGA 6054

Qy          1980 uProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLy 2000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6055 GCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAA 6114

Qy          2000 sAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAs 2020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6115 GAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGA 6174

Qy          2020 pProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6175 CCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTC 6234

Qy          2040 rLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSe 2060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6235 TCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAG 6294

Qy          2060 rSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSe 2080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6295 CTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAG 6354

Qy          2080 rHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAl 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6355 CCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGC 6414

Qy          2100 aAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLe 2120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6415 CGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCT 6474

Qy          2120 uGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSe 2140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6475 CCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAG 6534

Qy          2140 rGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPr 2160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6535 TGAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCC 6594

Qy          2160 oAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPr 2180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6595 CGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACC 6654

Qy          2180 oSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSe 2200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6655 CAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAG 6714

Qy          2200 rGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAs 2220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6715 CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGA 6774

Qy          2220 pGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAl 2240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6775 CGGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGC 6834

Qy          2240 aValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLy 2260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6835 TGTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAA 6894

Qy          2260 sSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSe 2280
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6895 GTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTC 6954

Qy         2280 rAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAs 2300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6955 CGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAA 7014

Qy         2300 nGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleTh 2320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7015 TGAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCAC 7074

Qy         2320 rGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMe 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7075 CGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACAT 7134

Qy         2340 tGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSe 2360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7135 GGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTC 7194

Qy         2360 rProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAl 2380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7195 CCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGC 7254

Qy         2380 aMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGl 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7255 TATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGG 7314

Qy         2400 yGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGl 2420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7315 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGG 7374

Qy         2420 yLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAs 2440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7375 CCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAA 7434

Qy         2440 nArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerTh 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7435 CCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCAC 7494

Qy         2460 rProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPr 2480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7495 GCCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACC 7554

Qy         2480 oProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGl 2500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7555 CCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGA 7614

Qy         2500 uGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7615 GGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7666

RESULT 6
ADQ84524
ID   ADQ84524 standard; cDNA; 8533 BP.
XX
AC   ADQ84524;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Human tumour-associated antigenic target (TAT) cDNA sequence #1338.
XX
KW   human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW   cancer; cell proliferative disorder; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004060270-A2.
XX
PD   22-JUL-2004.
XX
PF   15-OCT-2003; 2003WO-US029126.
XX
PR   18-OCT-2002; 2002US-0418988P.
XX
PA   (GETH ) GENENTECH INC.
PA   (WUTD/) WU T D.
PA   (ZHOU/) ZHOU Y.
XX
PI   Wu TD,  Zhou Y;
XX
DR   WPI; 2004-534300/51.
XX
PT   New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT   preventing or treating cell proliferative disorders such as cancer.
XX
PS   Claim 1; SEQ ID NO 1338; 5504pp; English.
XX
CC   The present invention describes an isolated tumour-associated antigenic
CC   target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC   sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC   (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC   sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC   (c). Also described: (1) an expression vector comprising the above
CC   nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC   a process for producing a polypeptide; (4) an isolated polypeptide
CC   comprising: (a) an amino acid sequence encoded by any of the above
CC   nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC   length coding region of the above nucleotide sequences; or (c) a sequence
CC   having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC   comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC   an isolated antibody that binds to the above polypeptide; (7) a process
CC   for producing the antibody; (8) an isolated oligopeptide that binds to
CC   the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC   binding organic molecule that binds to the above polypeptide; (10) a
CC   composition of matter comprising the above (chimeric) polypeptide,
CC   antibody, oligopeptide or TAT binding organic molecule, in combination
CC   with a carrier; (11) an article of manufacture comprising a container and
CC   the composition of matter contained within the container; (12) methods of
CC   inhibiting the growth of a cell that expresses the above protein, where
CC   the growth of the cell is at least in part dependent upon a growth
CC   potentiating effect of the above protein; (13) a method of
CC   therapeutically treating a mammal having a cancerous tumour comprising
CC   cells that express the above protein; (14) a method of determining the
CC   presence of a protein in a sample suspected of containing the protein
CC   described above; (15) methods of diagnosing the presence of a tumour in a
CC   mammal; (16) a method for treating or preventing a cell proliferative
CC   disorder associated with increased expression or activity of the above
CC   protein; and (17) a method of binding an antibody, oligopeptide or
CC   organic molecule to a cell that expresses the protein described above.
CC   The TAT sequences have cytostatic activities, and can be used in gene
CC   therapy. The composition and methods are useful for diagnosing,
CC   preventing or treating cancer. The composition is also used for preparing
CC   a medicament for the therapeutic treatment or diagnostic detection of a
CC   cell proliferative disorder or cancer. The present sequence represents a
CC   human TAT cDNA sequence from the present invention.
XX
SQ   Sequence 8533 BP; 1836 A; 3034 C; 2528 G; 1135 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0             Length:        8533
Score:                  13119.50      Matches:        2506
Percent Similarity:     99.56%        Conservative:  1
Best Local Similarity:  99.52%        Mismatches:    7
Query Match:            99.28%        Indels:        5
DB:                     13            Gaps:          1

US-09-522-753-5 (1-2517) x ADQ84524 (1-8533)

Qy            1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db            2 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy           21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121
```

```
Qy        41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 181

Qy        61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy        81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy       101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy       121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy       141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy       161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy       181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy       201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy       221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy       241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy       261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy       281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
             ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 901

Qy       301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
             ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       902 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 961

Qy       321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       962 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1021

Qy       341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy       361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy       381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy       401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy       421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy       441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy       461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy       481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAGCAACAACAGCAG 1501

Qy       501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy       521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy       541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy       561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy       581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy       601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy       621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy       641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy       661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy       681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy       701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy       721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy       741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy       761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy    781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
          |||||||||||||||||| ||||||||||||||||   |||   ||||||||||||||||
Db   2342 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCGGCCTC 2400

Qy    800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCC 2460

Qy    820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2520

Qy    840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2580

Qy    860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2640

Qy    880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2700

Qy    900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTAC 2760

Qy    920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2820

Qy    940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2880

Qy    960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 2940

Qy    980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3000

Qy   1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3060

Qy   1020 gGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGl 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GGGCAAGAGCAGGAGCCCGGCACCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCA 3120

Qy   1040 nLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProAr 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCG 3180

Qy   1060 gGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPr 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 TGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACC 3240

Qy   1080 oGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPr 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 TGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACC 3300

Qy   1100 oThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluAr 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAG 3360

Qy   1120 gGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHi 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCA 3420

Qy   1140 sAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLy 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAA 3480

Qy   1160 sLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPr 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 GCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCC 3540

Qy   1180 oProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLe 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCT 3600

Qy   1200 uGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAs 1220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGA 3660

Qy   1220 pSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLy 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CAGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAA 3720

Qy   1240 sGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGl 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 GGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGA 3780

Qy   1260 uAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTy 1280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTA 3840

Qy   1280 rGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPr 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACC 3900

Qy   1300 oProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyAr 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 CCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAG 3960

Qy   1320 gAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHi 1340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 AGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACA 4020

Qy   1340 sSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePr 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CAGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCC 4080

Qy   1360 oArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysAr 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 TCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCG 4140

Qy   1380 gGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGl 1400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 GGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCA 4200

Qy   1400 nAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGl 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 GGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGA 4260

Qy   1420 uAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPr 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 GGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCC 4320

Qy   1440 oLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAs 1460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 CCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGA 4380

Qy   1460 pThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPr 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 CACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCC 4440

Qy   1480 oGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGl 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 CGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGA 4500
```

```
Qy       1500 uArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGl 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 ACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGG 4560

Qy       1520 ySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPr 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 CTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCC 4620

Qy       1540 oLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVa 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 CCTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGT 4680

Qy       1560 lThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSe 1580
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4681 GACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATC 4740

Qy       1580 rGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVa 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4741 CCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGT 4800

Qy       1600 lProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGl 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4801 GCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGG 4860

Qy       1620 yValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGl 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4861 CGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGG 4920

Qy       1640 yIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProTh 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4921 CATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCAC 4980

Qy       1660 rTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGl 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 CTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGA 5040

Qy       1680 uAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAl 1700
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db       5041 GAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGC 5100

Qy       1700 aThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLe 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5101 CACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCT 5160

Qy       1720 uAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLe 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5161 GGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCT 5220

Qy       1740 uProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5221 GCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTA 5280

Qy       1760 rLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGl 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5281 CCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGG 5340

Qy       1780 yGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspAr 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5341 AGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCG 5400

Qy       1800 gAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrTh 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGAC 5460

Qy       1820 rValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySe 1840
              ||||||||||||||||||||||||||||||||||||||||         |||||||||||
Db       5461 GGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAG---------AGCAGCGGCAG 5511

Qy       1840 rSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHi 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5512 CAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCA 5571

Qy       1860 sSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAs 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5572 CTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAA 5631

Qy       1880 nThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSe 1900
              ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db       5632 CACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGGTC 5691

Qy       1900 rThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPr 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5692 CACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCC 5751

Qy       1920 oLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLy 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5752 ACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAA 5811

Qy       1940 sGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLe 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5812 GGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCT 5871

Qy       1960 uAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGl 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5872 CGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGA 5931

Qy       1980 uProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLy 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5932 GCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAA 5991

Qy       2000 sAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAs 2020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5992 GAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGA 6051

Qy       2020 pProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6052 CCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTC 6111

Qy       2040 rLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSe 2060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6112 TCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAG 6171

Qy       2060 rSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSe 2080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6172 CTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAG 6231

Qy       2080 rHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAl 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6232 CCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGC 6291

Qy       2100 aAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLe 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6292 CGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCT 6351

Qy       2120 uGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSe 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6352 CCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAG 6411

Qy       2140 rGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPr 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6412 TGAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCC 6471

Qy       2160 oAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPr 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6472 CGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACC 6531

Qy       2180 oSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSe 2200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6532 CAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAG 6591

Qy       2200 rGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAs 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6592 CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGA 6651

Qy       2220 pGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAl 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6652 CGGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGC 6711

Qy       2240 aValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLy 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6712 TGTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAA 6771

Qy       2260 sSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSe 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6772 GTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTC 6831

Qy       2280 rAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAs 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6832 CGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAA 6891

Qy       2300 nGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleTh 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6892 TGAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCAC 6951

Qy       2320 rGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMe 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6952 CGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACAT 7011

Qy       2340 tGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSe 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7012 GGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTC 7071

Qy       2360 rProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAl 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7072 CCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGC 7131

Qy       2380 aMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGl 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7132 TATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGG 7191

Qy       2400 yGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGl 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7192 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGG· 7251

Qy       2420 yLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAs 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7252 CCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAA 7311

Qy       2440 nArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerTh 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7312 CCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCAC 7371

Qy       2460 rProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPr 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7372 GCCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACC 7431

Qy       2480 oProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGl 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7432 CCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGA 7491

Qy       2500 uGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7492 GGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7543

RESULT 7
ACN39603
ID   ACN39603 standard; cDNA; 8533 BP.
XX
AC   ACN39603;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Tumour-associated antigenic target (TAT) cDNA DNA325805, SEQ ID NO:3856.
XX
KW   Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW   tumour; diagnosis; cell proliferative disorder; breast cancer;
KW   colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW   central nervous system cancer; bladder cancer; pancreatic cancer;
KW   cervical cancer; melanoma; leukaemia; hybridisation probe;
KW   chromosome identification; chromosome mapping; gene mapping;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004030615-A2.
XX
PD   15-APR-2004.
XX
PF   29-SEP-2003; 2003WO-US028547.
XX
PR   02-OCT-2002; 2002US-0414971P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wu TD,  Zhang Z,  Zhou Y;
XX
DR   WPI; 2004-347921/32.
XX
PT   New tumor-associated antigenic target polypeptides and nucleic acids,
PT   useful in preparing a medicament for treating or detecting a
PT   proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT   prostate cancer or tumor.
XX
PS   Claim 1; SEQ ID NO 3856; 7273pp; English.
XX
CC   The invention relates to human tumour-associated antigenic target (TAT)
CC   polypeptides, and their related nucleic acids. The TAT polypeptides are
CC   overexpressed in cancer tissues compared to normal tissues, and may thus
CC   serve as effective targets for the diagnosis and treatment of cancer in
CC   mammals. The invention also relates to nucleic acid and polypeptide
CC   sequences at least 80% identical to the TAT nucleic acids and
CC   polypeptides; expression vectors and host cells comprising a TAT nucleic
CC   acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC   molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC   TAT polypeptide; and methods and compositions for the treatment or
CC   diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC   antibodies, antagonists, binding molecules and compositions are useful
CC   for diagnosing or treating a cell proliferative disorder associated with
CC   increased TAT expression, particularly cancers such as breast cancer,
CC   colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC   cancer, pancreatic cancer, cervical cancer, cancers of the central
CC   nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC   used as hybridisation probes, in chromosome and gene mapping, in
CC   chromosome identification and in gene therapy. The present sequence
CC   represents a TAT nucleic acid of the invention
XX
SQ   Sequence 8533 BP; 1836 A; 3034 C; 2528 G; 1135 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               0              Length:        8533
Score:                   13119.50        Matches:       2506
Percent Similarity:      99.56%         Conservative:  1
Best Local Similarity:   99.52%         Mismatches:    7
Query Match:             99.28%         Indels:        5
DB:                      13             Gaps:          1

US-09-522-753-5 (1-2517) x ACN39603 (1-8533)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db          2 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 181

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 901

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        902 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 961

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1021

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321
```

```
Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAGCAACAACAGCAG 1501

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||||||||||||||||
Db       2342 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCGGCCTC 2400
```

```
Qy          800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2401 TGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCC 2460

Qy          820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2461 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2520

Qy          840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2521 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2580

Qy          860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2640

Qy          880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 GGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2700

Qy          900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTAC 2760

Qy          920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2820

Qy          940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2880

Qy          960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 2940

Qy          980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3000

Qy         1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3060

Qy         1020 gGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGl 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCA 3120

Qy         1040 nLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProAr 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 GAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCG 3180

Qy         1060 gGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPr 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 TGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACC 3240

Qy         1080 oGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPr 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 TGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACC 3300

Qy         1100 oThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluAr 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3301 CACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAG 3360

Qy         1120 gGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHi 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3361 GCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCA 3420

Qy         1140 sAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLy 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3421 TGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAA 3480

Qy         1160 sLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPr 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3481 GCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCC 3540

Qy         1180 oProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLe 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3541 ACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCT 3600

Qy         1200 uGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAs 1220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3601 GGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGA 3660

Qy         1220 pSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLy 1240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3661 CAGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAA 3720

Qy         1240 sGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGl 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3721 GGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGA 3780

Qy         1260 uAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTy 1280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3781 GGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTA 3840

Qy         1280 rGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPr 1300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3841 TGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACC 3900

Qy         1300 oProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyAr 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3901 CCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAG 3960

Qy         1320 gAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHi 1340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3961 AGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACA 4020

Qy         1340 sSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePr 1360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4021 CAGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCC 4080

Qy         1360 oArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysAr 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4081 TCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCG 4140

Qy         1380 gGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGl 1400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4141 GGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCA 4200

Qy         1400 nAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGl 1420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4201 GGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGA 4260

Qy         1420 uAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPr 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4261 GGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCC 4320

Qy         1440 oLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAs 1460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4321 CCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGA 4380

Qy         1460 pThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPr 1480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4381 CACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCC 4440

Qy         1480 oGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGl 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4441 CGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGA 4500

Qy         1500 uArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGl 1520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4501 ACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGG 4560

Qy         1520 ySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPr 1540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4561 CTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCC 4620

Qy        1540 oLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVa 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4621 CCTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGT 4680

Qy        1560 lThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSe 1580
               ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4681 GACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATC 4740

Qy        1580 rGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVa 1600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4741 CCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGT 4800

Qy        1600 lProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGl 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4801 GCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGG 4860

Qy        1620 yValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGl 1640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4861 CGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGG 4920

Qy        1640 yIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProTh 1660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4921 CATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCAC 4980

Qy        1660 rTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGl 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4981 CTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGA 5040

Qy        1680 uAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAl 1700
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        5041 GAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGC 5100

Qy        1700 aThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLe 1720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5101 CACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCT 5160

Qy        1720 uAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLe 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5161 GGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCT 5220

Qy        1740 uProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5221 GCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTA 5280

Qy        1760 rLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGl 1780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5281 CCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGG 5340

Qy        1780 yGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspAr 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5341 AGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCG 5400

Qy        1800 gAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrTh 1820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGAC 5460

Qy        1820 rValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySe 1840
               ||||||||||||||||||||||||||||||||||||||||         |||||||||||
Db        5461 GGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAG---------AGCAGCGGCAG 5511

Qy        1840 rSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHi 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5512 CAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCA 5571

Qy        1860 sSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAs 1880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5572 CTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAA 5631

Qy        1880 nThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSe 1900
               ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db        5632 CACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGGTC 5691

Qy        1900 rThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPr 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5692 CACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCC 5751

Qy        1920 oLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLy 1940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5752 ACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAA 5811

Qy        1940 sGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLe 1960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5812 GGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCT 5871

Qy        1960 uAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGl 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5872 CGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGA 5931

Qy        1980 uProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLy 2000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5932 GCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAA 5991

Qy        2000 sAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAs 2020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5992 GAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGA 6051

Qy        2020 pProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6052 CCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTC 6111

Qy        2040 rLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSe 2060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6112 TCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAG 6171

Qy        2060 rSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSe 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6172 CTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAG 6231

Qy        2080 rHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAl 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6232 CCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGC 6291

Qy        2100 aAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLe 2120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6292 CGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCT 6351

Qy        2120 uGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSe 2140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6352 CCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAG 6411

Qy        2140 rGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPr 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6412 TGAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCC 6471

Qy        2160 oAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPr 2180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6472 CGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACC 6531

Qy        2180 oSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSe 2200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6532 CAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAG 6591

Qy        2200 rGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAs 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6592 CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGA 6651

Qy        2220 pGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAl 2240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6652 CGGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGC 6711

Qy        2240 aValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLy 2260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6712 TGTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAA 6771
```

```
Qy       2260 sSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSe 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6772 GTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTC 6831

Qy       2280 rAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAs 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6832 CGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAA 6891

Qy       2300 nGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleTh 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6892 TGAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCAC 6951

Qy       2320 rGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMe 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6952 CGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACAT 7011

Qy       2340 tGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSe 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7012 GGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTC 7071

Qy       2360 rProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAl 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7072 CCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGC 7131

Qy       2380 aMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGl 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7132 TATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGG 7191

Qy       2400 yGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGl 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7192 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGG 7251

Qy       2420 yLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAs 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7252 CCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAA 7311

Qy       2440 nArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerTh 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7312 CCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCAC 7371

Qy       2460 rProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPr 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7372 GCCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACC 7431

Qy       2480 oProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGl 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7432 CCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGA 7491

Qy       2500 uGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7492 GGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7543

RESULT 8
ACA62250
ID   ACA62250 standard; cDNA; 7521 BP.
XX
AC   ACA62250;
XX
DT   12-AUG-2003  (first entry)
XX
DE   Human nuclear receptor corepressor SMRTe coding region cDNA.
XX
KW   Human; ss; gene; SMRTe; nuclear receptor corepressor; gene therapy;
KW   tissue typing; cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .7524
FT                   /*tag=  a
FT                   /partial
FT                   /product= "SMRTe"
FT                   /note= "No stop codon given"
XX
PN   US2003027137-A1.
XX
PD   06-FEB-2003.
XX
PF   27-MAR-2001; 2001US-00819104.
XX
PR   29-MAR-2000; 2000US-0193138P.
XX
PA   (CHEN/) CHEN J D.
XX
PI   Chen JD;
XX
DR   WPI; 2003-466139/44.
DR   P-PSDB; ABU61812.
XX
PT   New SMRTe proteins and nucleic acids, useful in gene therapy, predictive
PT   medicine, therapeutic or prophylactic treatment, chromosome mapping,
PT   tissue typing and in forensic biology.
XX
PS   Claim 2; Page 48-56; 90pp; English.
XX
CC   The invention relates to an isolated SMRTe nucleic acid molecule. The
CC   nucleic acids are useful in gene therapy, as hybridisation probes for
CC   identifying SMRTe-encoding nucleic acid molecules and as primers for
CC   amplifying of SMRTe nucleic acid molecules. The polypeptides are useful
CC   as immunogens to raise anti-SMRTe antibodies. The SMRTe molecules are
CC   useful as targets for discovering and developing modulating agents to
CC   regulate a variety of cellular processes, in screening assays, in
CC   predictive medicine, in therapeutic or prophylactic treatment, in
CC   chromosome mapping, tissue typing and in forensic identification of a
CC   biological sample. Modulators of SMRTe are useful for treating or
CC   preventing a condition associated with aberrant SMRTe protein or nucleic
CC   acid expression or activity, such as cancer. The present sequence
CC   represents the human nuclear receptor corepressor SMRTe coding region
CC   cDNA
XX
SQ   Sequence 7521 BP; 1635 A; 2728 C; 2212 G; 946 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0              Length:        7521
Score:                  12978.00        Matches:       2487
Percent Similarity:     98.57%         Conservative:  3
Best Local Similarity:  98.46%         Mismatches:    9
Query Match:            98.21%         Indels:        28
DB:                     8              Gaps:          3

US-09-522-753-5 (1-2517) x ACA62250 (1-7521)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360
```

```
Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 900

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        901 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1020

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1077

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1137

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1197

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1257

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1317

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1377

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1437

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1497

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1557

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1617

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1677

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1737

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1797

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1857

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1858 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1917

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1918 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1977

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1978 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2037

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2038 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2097

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db       2098 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2157

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||                                                   |||
Db       2158 GCTGAA---------------------------------------------------GCC 2166

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2226

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2286

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db       2287 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2345

Qy        800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
              ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db       2346 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2405

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2406 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2465

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2466 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2525

Qy         860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2526 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2585

Qy         880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
               |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        2586 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2645

Qy         900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2646 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2705

Qy         920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2706 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2765

Qy         940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2766 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2825

Qy         960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2826 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 2885

Qy         980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2886 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 2945

Qy        1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2946 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3005

Qy        1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
               ||||||||||||||||||||||||||||||||||||||||
Db        3006 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3065

Qy        1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3066 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3125

Qy        1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3126 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3185

Qy        1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3186 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3245

Qy        1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3246 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3305

Qy        1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3306 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3365

Qy        1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3366 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3425

Qy        1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3426 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3485

Qy        1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3486 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3545

Qy        1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3546 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3605

Qy        1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3606 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3665

Qy        1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3666 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3725

Qy        1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3726 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3785

Qy        1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3786 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 3845

Qy        1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3846 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 3905

Qy        1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3906 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 3965

Qy        1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3966 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4025

Qy        1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4026 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4085

Qy        1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4086 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4145

Qy        1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4146 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4205

Qy        1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4206 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4265

Qy        1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4266 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4325

Qy        1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4326 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4385

Qy        1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4386 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4445

Qy        1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4446 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4505

Qy        1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4506 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4565

Qy        1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4566 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4625

Qy        1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4626 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4685
```

```
Qy        1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4686 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4745

Qy        1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4746 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4805

Qy        1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4806 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 4865

Qy        1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4866 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 4925

Qy        1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4926 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 4985

Qy        1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4986 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5045

Qy        1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5046 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5105

Qy        1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5106 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5165

Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5166 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5225

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5226 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5285

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5286 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5345

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5346 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5405

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5406 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5465

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5466 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5525

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5526 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5585

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5586 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5645

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5646 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5705

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5706 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5765

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5766 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5825

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5826 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 5885

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5886 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 5945

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5946 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6005

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6006 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6065

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6066 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6125

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6126 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6185

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6186 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6245

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6246 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6305

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6306 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6365

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6366 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6425

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6426 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6485

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6486 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6545

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6546 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6605

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6606 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6665

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6666 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6725

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6726 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6785

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6786 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 6845

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6846 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 6905

Qy       2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6906 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 6965

Qy       2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6966 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7025

Qy       2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7026 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7085

Qy       2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7086 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7145

Qy       2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7146 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7205

Qy       2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7206 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7265

Qy       2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7266 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7325

Qy       2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7326 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7385

Qy       2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7386 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGCCCCCTCGC 7445

Qy       2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7446 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7505

Qy       2512 rLeuSerAspSerGlu 2517
              ||||||||||||||||
Db       7506 ACTCTCCGACAGCGAG 7521

RESULT 9
ADL13812
ID   ADL13812 standard; DNA; 7524 BP.
XX
AC   ADL13812;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Osteoarthritis-associated polymorphic nucleotide #344.
XX
KW   ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW   joint space narrowing; osteophyte development; joint pain;
KW   osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO2003054166-A2.
XX
PD   03-JUL-2003.
XX
PF   19-DEC-2002; 2002WO-US041225.
XX
PR   20-DEC-2001; 2001US-0342603P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Jones KA,  Schafer A;
```

```
XX
DR   WPI; 2003-559141/52.
XX
PT   Determining susceptibility of an individual to joint space narrowing,
PT   osteophyte development and/or joint pain comprises identifying whether
PT   the individual has at least one polymorphism in a polynucleotide encoding
PT   a protein.
XX
PS   Disclosure; SEQ ID NO 344; 297pp; English.
XX
CC   The invention relates to a method of determining susceptibility of an
CC   individual to joint space narrowing and/or osteophyte development and/or
CC   joint pain comprising identifying whether the individual has at least one
CC   polymorphism in a polynucleotide encoding at least one of the protein
CC   listed in the specification. The methods, composition and agent are
CC   useful for modulating the susceptibility of an individual to joint space
CC   narrowing and/or osteophyte development and/or joint pain that is
CC   associated with a disease, preferably osteoarthritis. The cell line and
CC   the non-human animal are useful for screening for an agent for diagnosing
CC   an individual having susceptibility to joint space narrowing and/or
CC   osteophyte development and/or joint pain. This sequence corresponds to
CC   the polynucleotide encoding a protein listed in the specification. (Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences).
XX
SQ   Sequence 7524 BP; 1636 A; 2728 C; 2213 G; 947 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0              Length:        7524
Score:                  12978.00        Matches:       2487
Percent Similarity:     98.57%         Conservative:  3
Best Local Similarity:  98.46%         Mismatches:    9
Query Match:            98.21%         Indels:        28
DB:                     10             Gaps:          3

US-09-522-753-5 (1-2517) x ADL13812 (1-7524)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540
```

```
Qy          181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy          201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 900

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db          901 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          961 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1020

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGCCAGCGGGCAGTGGG 1077

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         1078 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1137

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1138 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1197

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1198 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1257

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1258 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1317

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1318 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1377

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1378 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1437

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1438 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1497

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1557

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1558 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1617

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1618 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1677

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1678 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1737

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1738 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1797

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1798 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1857

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1858 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1917

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1918 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1977

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1978 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2037

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2038 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2097

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         2098 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGAAATGAGGAGGAGATGGTGGAGGAG 2157

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||                                                   |||
Db         2158 GCTGAA---------------------------------------------------GCC 2166

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2167 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2226

Qy          761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2227 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2286

Qy          781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
                |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db         2287 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2345

Qy          800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
                ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db         2346 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2405

Qy          820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2406 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2465

Qy          840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2466 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2525

Qy          860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2526 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2585

Qy          880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
                |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         2586 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2645

Qy          900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2646 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCAGGACAGCGACTCCAGTGCCAC 2705

Qy          920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2706 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2765

Qy          940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2766 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2825

Qy          960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2826 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 2885

Qy          980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2886 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 2945

Qy         1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2946 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3005

Qy         1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
                ||||||||||||||||||||||||||||||||||||||||
Db         3006 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3065

Qy         1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3066 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3125

Qy         1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3126 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3185

Qy         1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3186 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3245

Qy         1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3246 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3305

Qy         1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3306 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3365

Qy         1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3366 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3425

Qy         1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3426 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3485

Qy         1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3486 GTCCCCACGGGGCCAGGCTGGCCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3545

Qy         1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3546 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3605

Qy         1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3606 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3665

Qy         1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3666 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3725

Qy         1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3726 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3785

Qy         1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3786 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 3845

Qy         1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3846 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 3905

Qy         1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3906 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 3965

Qy         1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3966 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4025

Qy         1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4026 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4085

Qy         1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4086 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4145

Qy         1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4146 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4205

Qy         1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4206 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4265

Qy         1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4266 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4325

Qy         1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4326 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4385

Qy         1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4386 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4445

Qy         1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4446 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4505

Qy         1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4506 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4565

Qy         1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4566 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4625

Qy         1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4626 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4685

Qy         1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4686 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4745

Qy         1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4746 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4805

Qy         1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4806 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 4865
```

```
Qy        1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4866 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 4925

Qy        1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4926 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 4985

Qy        1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4986 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5045

Qy        1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5046 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5105

Qy        1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5106 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5165

Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5166 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5225

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5226 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5285

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5286 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5345

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5346 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5405

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5406 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5465

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5466 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5525

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5526 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5585

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5586 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5645

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5646 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5705

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5706 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5765

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5766 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5825

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5826 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 5885

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5886 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 5945

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5946 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6005

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6006 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6065

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6066 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6125

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6126 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6185

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6186 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6245

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6246 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6305

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6306 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6365

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6366 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6425

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6426 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6485

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6486 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6545

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6546 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6605

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6606 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6665

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6666 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6725

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6726 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6785

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6786 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 6845

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6846 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 6905

Qy        2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6906 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 6965

Qy        2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6966 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7025

Qy        2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7026 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7085

Qy        2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7086 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7145

Qy        2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7146 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7205

Qy        2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7206 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7265

Qy        2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7266 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7325

Qy        2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7326 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7385

Qy        2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7386 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7445

Qy        2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7446 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7505

Qy        2512 rLeuSerAspSerGlu 2517
               ||||||||||||||||
Db        7506 ACTCTCCGACAGCGAG 7521

RESULT 10
ACA62249
ID   ACA62249 standard; cDNA; 8686 BP.
XX
AC   ACA62249;
XX
DT   12-AUG-2003  (first entry)
XX
DE   cDNA encoding human nuclear receptor corepressor SMRTe.
XX
KW   Human; ss; gene; SMRTe; nuclear receptor corepressor; gene therapy;
KW   tissue typing; cancer.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             157. .7680
FT                   /*tag=  a
FT                   /product= "SMRTe"
XX
PN   US2003027137-A1.
XX
PD   06-FEB-2003.
XX
PF   27-MAR-2001; 2001US-00819104.
XX
PR   29-MAR-2000; 2000US-0193138P.
XX
PA   (CHEN/) CHEN J D.
XX
PI   Chen JD;
XX
DR   WPI; 2003-466139/44.
DR   P-PSDB; ABU61812.
XX
PT   New SMRTe proteins and nucleic acids, useful in gene therapy, predictive
PT   medicine, therapeutic or prophylactic treatment, chromosome mapping,
PT   tissue typing and in forensic biology.
```

```
XX
PS   Claim 2; Page 32-41; 90pp; English.
XX
CC   The invention relates to an isolated SMRTe nucleic acid molecule. The
CC   nucleic acids are useful in gene therapy, as hybridisation probes for
CC   identifying SMRTe-encoding nucleic acid molecules and as primers for
CC   amplifying of SMRTe nucleic acid molecules. The polypeptides are useful
CC   as immunogens to raise anti-SMRTe antibodies. The SMRTe molecules are
CC   useful as targets for discovering and developing modulating agents to
CC   regulate a variety of cellular processes, in screening assays, in
CC   predictive medicine, in therapeutic or prophylactic treatment, in
CC   chromosome mapping, tissue typing and in forensic identification of a
CC   biological sample. Modulators of SMRTe are useful for treating or
CC   preventing a condition associated with aberrant SMRTe protein or nucleic
CC   acid expression or activity, such as cancer. The present sequence
CC   represents cDNA encoding the human nuclear receptor corepressor SMRTe
XX
SQ   Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                0               Length:         8686
Score:                    12978.00         Matches:        2487
Percent Similarity:       98.57%          Conservative:   3
Best Local Similarity:    98.46%          Mismatches:     9
Query Match:              98.21%          Indels:         28
DB:                       8               Gaps:           3

US-09-522-753-5 (1-2517) x ACA62249 (1-8686)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db         157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy         161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy         181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy         201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816
```

```
Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1056

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db         1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||                                                   |||
Db         2314 GCTGAA---------------------------------------------------GCC 2322

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy          761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2383 GACACAGGGCAGAATGGCCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy          781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
                |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db         2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy          800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
                ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db         2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy          820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy          840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy          860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy          880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
                |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy          900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy          920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy          940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy          960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy          980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy         1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy         1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
                ||||||||||||||||||||||||||||||||||||||||
Db         3162 GGGCAAGAGCAGGAGCCCGGCACCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy         1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy         1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy         1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy         1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy         1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy         1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy         1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy         1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy         1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy         1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy         1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy         1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy         1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy         1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy         1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy         1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy         1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy         1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy         1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy         1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy         1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy         1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy         1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy         1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy         1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy         1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy         1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy         1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy         1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy         1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy         1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy         1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141
```

```
Qy        1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy        1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy        1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy        2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy        2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy        2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy        2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy        2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy        2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy        2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy        2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy        2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy        2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy        2512 rLeuSerAspSerGlu 2517
               ||||||||||||||||
Db        7662 ACTCTCCGACAGCGAG 7677

RESULT 11
ADL13811
ID   ADL13811 standard; DNA; 8686 BP.
XX
AC   ADL13811;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Osteoarthritis-associated polymorphic nucleotide #343.
XX
KW   ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW   joint space narrowing; osteophyte development; joint pain;
KW   osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO2003054166-A2.
XX
PD   03-JUL-2003.
XX
PF   19-DEC-2002; 2002WO-US041225.
XX
PR   20-DEC-2001; 2001US-0342603P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Jones KA,  Schafer A;
XX
DR   WPI; 2003-559141/52.
XX
PT   Determining susceptibility of an individual to joint space narrowing,
PT   osteophyte development and/or joint pain comprises identifying whether
PT   the individual has at least one polymorphism in a polynucleotide encoding
PT   a protein.
XX
PS   Disclosure; SEQ ID NO 343; 297pp; English.
XX
CC   The invention relates to a method of determining susceptibility of an
CC   individual to joint space narrowing and/or osteophyte development and/or
CC   joint pain comprising identifying whether the individual has at least one
CC   polymorphism in a polynucleotide encoding at least one of the protein
CC   listed in the specification. The methods, composition and agent are
CC   useful for modulating the susceptibility of an individual to joint space
CC   narrowing and/or osteophyte development and/or joint pain that is
CC   associated with a disease, preferably osteoarthritis. The cell line and
CC   the non-human animal are useful for screening for an agent for diagnosing
CC   an individual having susceptibility to joint space narrowing and/or
```

```
CC   osteophyte development and/or joint pain. This sequence corresponds to
CC   the polynucleotide encoding a protein listed in the specification. (Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences).
XX
SQ   Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0               Length:         8686
Score:                  12978.00         Matches:        2487
Percent Similarity:     98.57%          Conservative:   3
Best Local Similarity:  98.46%          Mismatches:     9
Query Match:            98.21%          Indels:         28
DB:                     10              Gaps:           3

US-09-522-753-5 (1-2517) x ADL13811 (1-8686)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db         157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy         161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy         181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy         201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy         221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy         241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy         261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996
```

```
Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1056

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db       2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||                                                   |||
Db       2314 GCTGAA---------------------------------------------------GCC 2322

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db       2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy        800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
              ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db       2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy        860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy        880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy        900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy        920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy        940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy        980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy       1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy         1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu------------------ 1033
                ||||||||||||||||||||||||||||||||||||||||||
Db         3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy         1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy         1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy         1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy         1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy         1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy         1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy         1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy         1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy         1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy         1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy         1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy         1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy         1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy         1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy         1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy         1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy         1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy         1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy         1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy         1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy         1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy         1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy         1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy         1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy         1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy         1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy         1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy         1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy         1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy         1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy         1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy         1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy         1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy         1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy         1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321
```

```
Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy        2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy        2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy        2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy        2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy        2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy        2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy        2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy        2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy       2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGCCCCTCGC 7601

Qy       2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy       2512 rLeuSerAspSerGlu 2517
              ||||||||||||||||
Db       7662 ACTCTCCGACAGCGAG 7677

RESULT 12
ADG86290
ID   ADG86290 standard; cDNA; 8686 BP.
XX
AC   ADG86290;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Human SMRT encoding cDNA SEQ ID NO:4.
XX
KW   SMRT; silencing mediator for retinoid and thyroid hormone action;
KW   SMRT inhibitor; cytostatic; antiinflammatory; antiarthritic;
KW   antirheumatic; antisense therapy; inflammatory disorder;
KW   rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;
KW   breast cancer; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             157. .7680
FT                   /*tag=  a
FT                   /product= "SMRT"
XX
PN   WO2003106645-A2.
XX
PD   24-DEC-2003.
XX
PF   17-JUN-2003; 2003WO-US018923.
XX
PR   17-JUN-2002; 2002US-00174014.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Bennett CF,  Freier SM,  Dobie KW;
XX
DR   WPI; 2004-082184/08.
DR   P-PSDB; ADG86291.
DR   GENBANK; AF125672.
XX
PT   Novel antisense compound targeted to nucleic acid encoding SMRT
PT   (silencing mediator for retinoid and thyroid hormone action), useful for
PT   treating animal having disease associated with SMRT such as cancer,
PT   rheumatoid arthritis.
XX
PS   Example 13; SEQ ID NO 4; 260pp; English.
XX
CC   The present invention describes a compound (I) 8-50 nucleobases in length
CC   targeted to a nucleic acid molecule encoding SMRT (silencing mediator for
CC   retinoid and thyroid hormone action), where (I) specifically hybridises
CC   with the nucleic acid molecule encoding SMRT and inhibits expression of
CC   SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a
CC   preferred target region on nucleic acid molecule encoding SMRT. Also
CC   described is a composition (II) comprising (I) and a carrier or diluent.
CC   (I) and (II) have cytostatic, antiinflammatory, antiarthritic and
CC   antirheumatic activities, and can be used in antisense therapy, and as
CC   SMRT expression inhibitors. (I) is useful for inhibiting the expression
CC   of SMRT in cells or tissues. (I) is also useful for treating an animal
CC   having a disease or condition associated with SMRT, e.g., inflammatory
CC   disorder such as rheumatoid arthritis; or a hyperproliferative disorder
CC   such as cancer chosen from leukaemia and breast cancer, by inhibiting the
CC   expression of SMRT. (I) is useful for diagnostics, therapeutics,
CC   prophylaxis and as research reagents and kits. The present sequence
CC   encodes human SMRT, which is used in an example from the present
CC   invention.
XX
SQ   Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0              Length:        8686
Score:                  12978.00        Matches:       2487
Percent Similarity:     98.57%         Conservative:  3
Best Local Similarity:  98.46%         Mismatches:    9
Query Match:            98.21%         Indels:        28
DB:                     12             Gaps:          3

US-09-522-753-5 (1-2517) x ADG86290 (1-8686)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1056

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db         1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||                                                   |||
Db         2314 GCTGAA---------------------------------------------------GCC 2322

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy          761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2383 GACACAGGGCAGAATGGCCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGCCCACCCCCA 2442

Qy          781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
                |||||||||||||||||| ||||||||||||||||  |||   |||   |||||||||||
Db         2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy          800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
                ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db         2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy          820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy          840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy          860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy          880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
                |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy          900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy          920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy          940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy          960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy          980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101
```

```
Qy     1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy     1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
            ||||||||||||||||||||||||||||||||||||||||
Db     3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy     1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy     1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy     1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy     1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy     1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy     1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy     1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy     1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy     1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy     1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy     1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy     1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy     1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy     1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy     1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy     1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy     1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy     1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy     1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy     1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy     1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy     1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy     1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy     1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy     1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy     1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy     1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy     1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy     1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy     1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy     1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy     1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy     1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy     1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy     1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy        2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy        2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy        2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy        2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy        2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy        2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy        2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481
```

```
Qy       2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy       2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy       2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy       2512 rLeuSerAspSerGlu 2517
              ||||||||||||||||
Db       7662 ACTCTCCGACAGCGAG 7677

RESULT 13
ADQ18920
ID   ADQ18920 standard; DNA; 8686 BP.
XX
AC   ADQ18920;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 1739.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 1739; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0              Length:        8686
Score:                  12978.00        Matches:       2487
Percent Similarity:     98.57%         Conservative:  3
Best Local Similarity:  98.46%         Mismatches:    9
```

```
Query Match:            98.21%         Indels:        28
DB:                     12             Gaps:          3

US-09-522-753-5 (1-2517) x ADQ18920 (1-8686)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1056

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176
```

```
Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db       2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||                                                   |||
Db       2314 GCTGAA---------------------------------------------------GCC 2322

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db       2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy        800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
              ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db       2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy        860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy        880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy        900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy        920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy        940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy        980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy       1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy       1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
              ||||||||||||||||||||||||||||||||||||||||
Db       3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy       1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy       1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy       1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy       1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy       1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy       1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy       1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy       1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy       1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy       1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy       1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy       1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy       1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy       1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy       1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy       1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy       1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy       1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy       1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy       1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy       1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy       1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy       1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy       1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy       1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy       1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy       1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy       1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy       1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy       1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy       1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy       1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501
```

```
Qy       1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy       1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy       1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy       1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy       1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy       1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy       1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy       1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy       1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy       1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy       1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy       2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy       2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy       2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy       2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy       2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy       2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy       2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy       2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy       2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy       2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy       2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy       2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy       2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy       2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy       2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy       2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy       2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy       2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy       2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy       2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy       2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy       2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy       2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy       2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy       2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy       2512 rLeuSerAspSerGlu 2517
              ||||||||||||||||
Db       7662 ACTCTCCGACAGCGAG 7677

RESULT 14
ADL12577
ID   ADL12577 standard; cDNA; 9053 BP.
XX
AC   ADL12577;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human steroid-induced C3A liver cell cDNA #306.
XX
KW   ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW   steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS   Homo sapiens.
XX
PN   US6673549-B1.
XX
PD   06-JAN-2004.
XX
PF   12-OCT-2001; 2001US-00976594.
XX
PR   12-OCT-2000; 2000US-0240409P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Furness LM,  Buchbinder JL;
XX
DR   WPI; 2004-068610/07.
XX
PT   Combination useful for preparing a composition for treating liver
PT   disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT   comprises cDNAs that are differentially expressed in response to steroid
PT   treatment.
XX
PS   Claim 1; SEQ ID NO 306; 141pp; English.
XX
CC   The invention relates to a combination comprising cDNAs that are
CC   differentially expressed in response to steroid treatment. Also included
CC   are the following: a high throughput method for using a cDNA to detect
CC   differential expression of nucleic acids in a sample; and a high
CC   throughput method of screening molecules or compounds to identify a
CC   ligand that specifically binds a cDNA. The sample is from a subject with
CC   Wilson disease and comparison of a standard defines a stage of that
CC   disease. The high throughput method of screening molecules or compounds
CC   to identify a ligand that specifically binds a cDNA comprises: combining
CC   the combination with molecules or compounds under conditions to allow
CC   specific binding; and detecting specific binding between each cDNA and at
CC   least one molecule or compound. The molecules or compounds are regulatory
CC   proteins. The combination is useful for preparing a composition for
CC   treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC   or hepatitis. The present sequence represents a human cDNA which is
CC   differentially expressed in steroid-induced C3A liver cells. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 9053 BP; 1904 A; 3171 C; 2722 G; 1254 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.:              0              Length:         9053
Score:                  12922.00        Matches:        2487
Percent Similarity:     98.26%         Conservative:   2
Best Local Similarity:  98.18%         Mismatches:     10
Query Match:            97.78%         Indels:         35
DB:                     12             Gaps:           5

US-09-522-753-5 (1-2517) x ADL12577 (1-9053)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        520 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 579

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 639

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 699

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 759

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 819

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 879

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 939

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 999

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 1059

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 1119

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 1179

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 1239

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 1299

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 1359

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       1360 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1419

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1420 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1479

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1480 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1539

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1599

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1600 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1659
```

```
Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||   :::||||||||||||||||||||||||||||||||||||||||||
Db       1660 GAGCAGGAGAACCGGAAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1719

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1720 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1779

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1780 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1839

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1840 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1899

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1959

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGln---GlnGlnGlnGln 499
              |||||||||||||||||||||||||||||||||||||||||||||   |||   ||||||
Db       1960 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCNACAACNGCAGCAG 2019

Qy        500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAG 2079

Qy        520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AAAGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAG 2139

Qy        540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 AACGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGAC 2199

Qy        560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 GAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAA 2259

Qy        580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 GGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAG 2319

Qy        600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2320 CAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAA 2379

Qy        620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2380 ATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGG 2439

Qy        640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2440 ATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2499

Qy        660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 CAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCG 2559

Qy        680 ArgArgLysLysLysLysAlaProAlaAlaAla-SerGluGluAlaAlaPheProProVa 699
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2560 CGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCCAGCGAGGAGGCTGCATTCCCGCCCGT 2619

Qy        699 lValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGl 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 GGTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGA 2679

Qy        719 uGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPr 739
              ||||||||||
Db       2680 GGAGGCTGAA-------------------------------------------------- 2689

Qy        739 oAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAl 759
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2690 -GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGC 2748

Qy        759 aLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAsp-GlyProP 779
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       2749 CAAGGACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGGCCAC 2808

Qy        779 roProGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrPro 798
              ||||||||||||||||||||||| |||||||||||||||   |||   ||||||||||||
Db       2809 CCCCAGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCG 2867

Qy        799 AlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaPro 818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2868 GCCTCTGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCT 2927

Qy        819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGlu 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2928 CCTCCTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAG 2987

Qy        839 GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2988 GAGGGGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCC 3047

Qy        859 GluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLys 878
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3048 GAGGAGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAG 3107

Qy        879 AspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGly 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3108 GACGCGGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGC 3167

Qy        899 GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3168 GGGAGCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGT 3227

Qy        919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3228 GCTACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTG 3287

Qy        939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3288 TCCCCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAG 3347

Qy        959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       3348 AAGCCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATC---GTC 3404

Qy        979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProPro-Al 998
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3405 ACCAAAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGAC 3464

Qy        998 aProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSe 1018
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3465 CCCACCGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAG 3524

Qy       1018 rProArgGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------- 1033
              ||||||||||||||||||||||||||||||||||||||||||||||
Db       3525 CCCCCGGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGT 3584

Qy       1034 ----------AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpTh 1050
                        |||||||||||||||||||||||||||||||||||.||||||||||||||
Db       3585 GTTCTTCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGAC 3644

Qy       1050 rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3645 TTCCGGCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCC 3704

Qy       1070 oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3705 GGACCCCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGA 3764

Qy       1090 pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3765 CACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTC 3824

Qy        1110 rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3825 CTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTC 3884

Qy        1130 rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3885 GGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCAT 3944

Qy        1150 tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3945 GGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGA 4004

Qy        1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4005 GCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCA 4064

Qy        1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4065 GGAGGCGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAA 4124

Qy        1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4125 AGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCAC 4184

Qy        1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4185 CCACGGCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGA 4244

Qy        1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4245 CAGCCCGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTA 4304

Qy        1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4305 CGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTC 4364

Qy        1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4365 CAAGGAGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCAC 4424

Qy        1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4425 CTATGACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCT 4484

Qy        1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4485 CATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCA 4544

Qy        1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4545 CATCCGCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTA 4604

Qy        1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4605 CCTGCGTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTC 4664

Qy        1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4665 ACGGGACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGC 4724

Qy        1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4725 CCATGAGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCG 4784

Qy        1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySe 1450
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4785 CGAGGAGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTC 4844

Qy        1450 rIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLy 1470
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4845 CATCACGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAA 4904

Qy        1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLe 1490
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4905 GCACGACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCT 4964

Qy        1490 uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSe 1510
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4965 GGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAG 5024

Qy        1510 rArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa 1530
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5025 CCGGCCAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGT 5084

Qy        1530 lProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPh 1550
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5085 GCCTGAGCTGGGCAAGCCGCGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTT 5144

Qy        1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl 1570
               ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db        5145 TGCCGGCCACCTCCCACGAGGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCA 5204

Qy        1570 nGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5205 GGAGGGCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCG 5264

Qy        1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5265 TGAGATCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCC 5324

Qy        1610 oTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLe 1630
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5325 CTATGAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCT 5384

Qy        1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTy 1650
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5385 GGCCTTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTA 5444

Qy        1650 rLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIl 1670
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5445 CCTGCCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCAT 5504

Qy        1670 eArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIl 1690
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5505 CCGCGGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACAT 5564

Qy        1690 eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
               |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        5565 CACCTCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCT 5624

Qy        1710 uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGl 1730
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5625 GAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGG 5684

Qy        1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5685 CATCATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCAC 5744

Qy        1750 rProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSe 1770
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5745 CCCAGCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAG 5804

Qy        1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh 1790
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5805 CCGCCACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCAC 5864

Qy        1790 rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr 1810
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5865 CACGTCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCG 5924

Qy        1810 gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGl 1830
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5925 GGAAAAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGG 5984
```

```
Qy        1830 yThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerAr 1850
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5985 TACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCG 6044

Qy        1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6045 CCCCGCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGC 6104

Qy        1870 aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa 1890
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6105 CCTCCAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGT 6164

Qy        1890 lGluProSerLysProThrValLeu---------ArgSerThrSerThrSerSerProVa 1907
               ||||||||||   ||||||||||||         ||||||||||||||||||||||||||
Db        6165 GGAGCCCAGCACGCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCCTCACCCGT 6224

Qy        1907 lArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGl 1927
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6225 TCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGG 6284

Qy        1927 yValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaAr 1947
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6285 GGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCG 6344

Qy        1947 gProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSe 1967
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6345 GCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTC 6404

Qy        1967 rGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProPr 1987
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6405 CGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCC 6464

Qy        1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAl 2007
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6465 TGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGC 6524

Qy        2007 aSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGl 2027
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6525 CAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCA 6584

Qy        2027 nSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSe 2047
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6585 AAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAG 6644

Qy        2047 rTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAs 2067
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6645 CTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGA 6704

Qy        2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuAr 2087
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6705 CAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCG 6764

Qy        2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6765 GCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCG 6824

Qy        2107 gProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLy 2127
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6825 GCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAA 6884

Qy        2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6885 AGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTA 6944

Qy        2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6945 CACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCC 7004

Qy        2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7005 TGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCC 7064

Qy        2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerPr 2207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7065 CCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCC 7124

Qy        2207 oGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerPr 2227
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7125 AGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCC 7184

Qy        2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7185 ACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCG 7244

Qy        2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGl 2267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7245 GGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCA 7304

Qy        2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLy 2287
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7305 GCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAA 7364

Qy        2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7365 GCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAG 7424

Qy        2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7425 CCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTA 7484

Qy        2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7485 TAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAG 7544

Qy        2347 gLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAl 2367
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7545 AAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGC 7604

Qy        2367 aPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAs 2387
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7605 TTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGA 7664

Qy        2387 pGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGl 2407
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7665 CGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGG 7724

Qy        2407 yArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPr 2427
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7725 CAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCC 7784

Qy        2427 oProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAs 2447
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7785 ACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAA 7844

Qy        2447 nArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLe 2467
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7845 CCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCT 7904

Qy        2467 uIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGl 2487
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7905 GATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGG 7964

Qy        2487 ySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCy 2507
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7965 CAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTG 8024

Qy        2507 sSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||
Db        8025 CTCGCAGTACGAGACACTCTCCGACAGCGAG 8055

RESULT 15
ADQ23294/c
ID   ADQ23294 standard; DNA; 9079 BP.
XX
AC   ADQ23294;
XX
```

```
DT   26-AUG-2004  (first entry)
XX
DE   Human soft tissue sarcoma-upregulated DNA - SEQ ID 6114.
XX
KW   soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO2004048938-A2.
XX
PD   10-JUN-2004.
XX
PF   26-NOV-2003; 2003WO-US038193.
XX
PR   26-NOV-2002; 2002US-0429739P.
XX
PA   (PROT-) PROTEIN DESIGN LABS INC.
XX
PI   Aziz N,  Ginsburg WM,  Zlotnik A;
XX
DR   WPI; 2004-441208/41.
XX
PT   Early detection of soft tissue sarcoma comprises determining expression
PT   of a gene in a first soft tissue sample and a normal soft tissue sample
PT   and comparing the gene expression, also useful in treating soft tissue
PT   sarcoma.
XX
PS   Example 2; SEQ ID NO 6114; 210pp; English.
XX
CC   The invention relates to a novel method for detecting soft tissue sarcoma
CC   which comprises obtaining a first soft tissue sample from an individual
CC   and a normal soft tissue sample from the same or different individual,
CC   determining the expression of a gene in both samples and comparing the
CC   expression of the gene in both soft tissue samples, where a higher level
CC   of protein expression in the first soft tissue sample indicates the
CC   presence of soft tissue sarcoma. The method of the invention has
CC   cytostatic applications and may be useful for detecting soft tissue
CC   sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC   acid sequences may be useful in diagnostic and screening applications.
CC   The current sequence is that of a human soft tissue sarcoma-upregulated
CC   DNA of the invention. The current sequence is not shown within the
CC   specification per se but was submitted in CD format by the inventor.
XX
SQ   Sequence 9079 BP; 1268 A; 2703 C; 3184 G; 1924 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              0            Length:         9079
Score:                  12845.50      Matches:        2473
Percent Similarity:     98.10%       Conservative:   6
Best Local Similarity:  97.86%       Mismatches:     20
Query Match:            97.20%       Indels:         31
DB:                     12           Gaps:           3

US-09-522-753-5 (1-2517) x ADQ23294 (1-9079)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       8552 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 8493

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8492 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 8433

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8432 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 8373

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8372 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 8313

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8312 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 8253

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8252 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 8193

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8192 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGGCC- 8134

Qy        141 LeuThrGlyLysLeuGluProValSer-ProProSerProProHisThrAspProGluLe 160
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       8133 CTGACGGGCAAGCTGGAACCGGTGTCTCCCACCCAGCCCCCCGCACACTGACCCTGAGCT 8074

Qy        160 uGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8073 GGAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGA 8014

Qy        180 pArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8013 CCGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCT 7954

Qy        200 uGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGl 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7953 GGAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAACCCGTGTCACCGCCGCCCATCGA 7894

Qy        220 uSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7893 GTCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGC 7834

Qy        240 aAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPr 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7833 TGCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCC 7774

Qy        260 oSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLe 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7773 CTCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCT 7714

Qy        280 uIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGl 300
              |||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db       7713 AATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCA 7654

Qy        300 nArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPr 320
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       7653 GCGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCC 7594

Qy        320 oArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIl 340
              ||||||||||||||||||||||||||||||||||||||||||||||.|||||||||||||
Db       7593 CCGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGAT 7534

Qy        340 eArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGl 360
              ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db       7533 CCGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGG 7477

Qy        360 yLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSe 380
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       7476 GCTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTC 7417

Qy        380 rGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTy 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7416 AGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTA 7357

Qy        400 rAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMe 420
              ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db       7356 CGACGCTGACCAGCAGCGCATCACGTTCATCAACATGAACGGGCTTATGGCCGACCCCAT 7297

Qy        420 tLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPh 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7296 GAAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTT 7237

Qy        440 eArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluAr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7236 CCGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAG 7177

Qy     460 gLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLy 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7176 GAAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAA 7117

Qy     480 sSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGl 500
           |||||||||||||||||||||||||||||||||||||:::   |||||||||   :::
Db    7116 GAGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAACC--CAGCAGCAAACAAAAGG 7059

Qy     500 nGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLy 520
            :::|||                  |||||||||||||||||||||||||||||||||||
Db    7058 CCGGCAGGCAGCAGCCACCAGCAGCCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAA 6999

Qy     520 sAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAs 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6998 AGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAA 6939

Qy     540 nAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGl 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6938 CGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGA 6879

Qy     560 uLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGl 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6878 GAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGG 6819

Qy     580 yArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGl 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6818 CCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCA 6759

Qy     600 nSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMe 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6758 GAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAAT 6699

Qy     620 tGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMe 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6698 GGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGAT 6639

Qy     640 tValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGl 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6638 GGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCA 6579

Qy     660 nAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaAr 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6578 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCG 6519

Qy     680 gArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVa 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6518 GAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGT 6459

Qy     700 lGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGl 720
           |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db    6458 GGAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGA 6399

Qy     720 uAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAl 740
           |||||||                                                   ||
Db    6398 GGCTGAA---------------------------------------------------GC 6390

Qy     740 aThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLy 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6389 CACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAA 6330

Qy     760 sAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPr 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6329 GGACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCC 6270

Qy     780 oGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaS 800
           ||||||||||||||||||| |||||||||||||||   |||   |||   ||||||||||
Db    6269 AGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCT 6211

Qy     800 erGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProP 820
           |||||||||||   |||||||||||||||||||||||||||   ||||||:::|||||||
Db    6210 CTGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTC 6151

Qy     820 roValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluG 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6150 CTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGG 6091

Qy     840 lyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluG 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6090 GGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGG 6031

Qy     860 luProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspA 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6030 AGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACG 5971

Qy     880 laGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlyS 900
           ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db    5970 CGGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGA 5911

Qy     900 erGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaT 920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5910 GCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCA 5851

Qy     920 hrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerP 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5850 CCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCC 5791

Qy     940 roArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysP 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5790 CAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGC 5731

Qy     960 roLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrL 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5730 CACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCA 5671

Qy     980 ysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProP 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5670 AAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCAC 5611

Qy    1000 roProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5610 CGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCC 5551

Qy    1020 rgGlyLysSerArgSerProAlaProProAlaAspLysGlu------------------- 1033
           |||||||||||||||||||||||||||||||||||||||||
Db    5550 GGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCT 5491

Qy    1034 -----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerG 1052
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5490 TCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCG 5431

Qy    1052 lyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspP 1072
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5430 GCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACC 5371

Qy    1072 roSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrA 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5370 CCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTG 5311

Qy    1092 laArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerA 1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5310 CCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTG 5251

Qy    1112 laLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValG 1132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5250 CCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCC 5191

Qy    1132 lnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyL 1152
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5190 AGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGC 5131

Qy    1152 euProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnL 1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5130 TGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGC 5071
```

```
Qy        1172 euSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluA 1192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5070 TGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGG 5011

Qy        1192 laSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyI 1212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5010 CGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCA 4951

Qy        1212 leProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisG 1232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4950 TTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACG 4891

Qy        1232 lyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerP 1252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4890 GCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCC 4831

Qy        1252 roSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluG 1272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4830 CGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAG 4771

Qy        1272 lyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysG 1292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4770 GCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGG 4711

Qy        1292 luAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrA 1312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4710 AGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATG 4651

Qy        1312 spMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetG 1332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4650 ACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGG 4591

Qy        1332 lyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleA 1352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4590 GCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCC 4531

Qy        1352 rgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuA 1372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4530 GCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGC 4471

Qy        1372 rgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgA 1392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4470 GTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGG 4411

Qy        1392 spLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisG 1412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4410 ACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATG 4351

Qy        1412 luGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluG 1432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4350 AGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGG 4291

Qy        1432 luLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleT 1452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4290 AGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCA 4231

Qy        1452 hrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisA 1472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4230 CGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACG 4171

Qy        1472 spValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspV 1492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4170 ACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATG 4111

Qy        1492 alMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgP 1512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4110 TGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGC 4051

Qy        1512 roGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProG 1532
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4050 CAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTG 3991

Qy        1532 luLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaG 1552
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3990 AGCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCG 3931

Qy        1552 lyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluG 1572
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3930 GCCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGG 3871

Qy        1572 lySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluI 1592
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3870 GCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGA 3811

Qy        1592 leAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrG 1612
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3810 TCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATG 3751

Qy        1612 luHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaP 1632
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3750 AGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCT 3691

Qy        1632 heAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuP 1652
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3690 TCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGC 3631

Qy        1652 roArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgG 1672
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3630 CCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCG 3571

Qy        1672 lyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrS 1692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3570 GCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCT 3511

Qy        1692 erGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgG 1712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3510 CGCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGG 3451

Qy        1712 lyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleI 1732
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3450 GCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCA 3391

Qy        1732 leAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProA 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3390 TCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAG 3331

Qy        1752 laThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgH 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3330 CCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCC 3271

Qy        1772 isSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrS 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3270 ACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGT 3211

Qy        1792 erSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluL 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3210 CCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAA 3151

Qy        1812 ysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrG 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3150 AGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAG 3091

Qy        1832 luGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProA 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3090 AGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCG 3031

Qy        1852 laSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuG 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3030 CCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCC 2971

Qy        1872 lnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluP 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2970 AGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGC 2911

Qy        1892 roSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaT 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2910 CCAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCA 2851

Qy        1912 hrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrL 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2850 CATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCC 2791

Qy        1932 euMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProA 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2790 TCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCC 2731

Qy        1952 rgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProA 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2730 GAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCG 2671

Qy        1972 laSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisA 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2670 CCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACG 2611

Qy        1992 laThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProP 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2610 CCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGC 2551

Qy        2012 roAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheS 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2550 CGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTT 2491

Qy        2032 erIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluG 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2490 CCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAG 2431

Qy        2052 lyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProL 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2430 GGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCA 2371

Qy        2072 ysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProG 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2370 AGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAG 2311

Qy        2092 lyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluS 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2310 GCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGA 2251

Qy        2112 erGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgV 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2250 GCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGG 2191

Qy        2132 alValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisP 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2190 TGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACC 2131

Qy        2152 roGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysP 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2130 CACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCC 2071

Qy        2172 roValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyA 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2070 CCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTG 2011

Qy        2192 laProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysT 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2010 CCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGA 1951

Qy        2212 hrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetT 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1950 CGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGA 1891

Qy        2232 hrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnT 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1890 CGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGA 1831

Qy        2252 hrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPheP 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1830 CGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCT 1771

Qy        2272 heSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnL 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1770 TCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACA 1711

Qy        2292 ysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrG 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1710 AGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGG 1651

Qy        2312 luIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaV 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1650 AGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGG 1591

Qy        2332 alGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetG 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1590 TGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGG 1531

Qy        2352 lyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuA 2372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1530 GTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGA 1471

Qy        2372 snAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspH 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1470 ATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACC 1411

Qy        2392 isThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerA 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1410 ACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCC 1351

Qy        2412 rgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerS 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1350 GAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCT 1291

Qy        2432 erValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluA 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1290 CAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGG 1231

Qy        2452 spArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuG 2472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1230 ACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGC 1171

Qy        2472 lnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuA 2492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1170 AGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCG 1111

Qy        2492 laGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluT 2512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1110 CTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGA 1051

Qy        2512 hrLeuSerAspSerGlu 2517
               |||||||||||||||||
Db        1050 CACTCTCCGACAGCGAG 1034

Search completed: November 11, 2005, 10:58:53
Job time : 4786 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         November 11, 2005, 09:39:08 ; Search time 5611 Seconds
                                           (without alignments)
                                           3709.701 Million cell updates/sec

Title:          US-09-522-753-5
Perfect score:  13215
Sequence:       1 MSGSTQLVAQTWRATEPRYP..........WDEEPKPLLCSQYETLSDSE 2517

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters:      19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09522753/runat_02112005_173641_20852/app_query.fasta_1.2695
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09522753_@CGN_1_1_1531_@runat_02112005_173641_20852
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
                14:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                19:  /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
                20:  /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
                21:  /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
                22:  /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
                23:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                24:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
                25:  /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
                26:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
                27:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                28:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

```
       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                         Description
---------------------------------------------------------------------------
   1    13215   100.0   7554  18  US-10-414-692-14           Sequence 14, Appl
   2    13215   100.0   8561  18  US-10-174-014-11           Sequence 11, Appl
   3    13215   100.0   8561  26  US-11-052-344-1            Sequence 1, Appli
   4    13145    99.5   8667  14  US-10-087-192-653          Sequence 653, App
   5    12978    98.2   7521  10  US-09-819-104A-3           Sequence 3, Appli
   6    12978    98.2   8686  10  US-09-819-104A-1           Sequence 1, Appli
   7    12978    98.2   8686  18  US-10-174-014-4            Sequence 4, Appli
   8    12978    98.2   8686  21  US-10-723-860-1739         Sequence 1739, Ap
c  9 12845.5     97.2   9079  21  US-10-723-860-6114         Sequence 6114, Ap
  10 10832.5     82.0   7386  10  US-09-819-104A-6           Sequence 6, Appli
  11 10832.5     82.0   8544  10  US-09-819-104A-4           Sequence 4, Appli
  12 10832.5     82.0   8544  22  US-10-764-420-1526         Sequence 1526, Ap
  13  9691.5     73.3   7534  14  US-10-087-192-650          Sequence 650, App
  14    4612     34.9   2930  16  US-10-146-473-14           Sequence 14, Appl
  15    4612     34.9   2930  18  US-10-174-014-13           Sequence 13, Appl
  16  4165.5     31.5   7940  18  US-10-341-434-235          Sequence 235, App
  17  4165.5     31.5   7940  22  US-10-887-553A-664         Sequence 664, App
  18  4147.5     31.4   7940  21  US-10-473-974-191          Sequence 191, App
  19  4147.5     31.4   7940  22  US-10-956-157-1580         Sequence 1580, Ap
  20  4147.5     31.4   7940  24  US-10-745-237-221          Sequence 221, App
  21    3649     27.6  91141  14  US-10-087-192-649          Sequence 649, App
  22    2951     22.3 233380  14  US-10-087-192-652          Sequence 652, App
  23    2184     16.5   2745  18  US-10-414-692-15           Sequence 15, Appl
  24  2133.5     16.1 221000  18  US-10-174-014-12           Sequence 12, Appl
  25    1083      8.2  10910  24  US-10-745-237-23           Sequence 23, Appl
  26    1083      8.2  10910  26  US-11-097-143-1937         Sequence 1937, Ap
  27    1076      8.1    872  24  US-10-779-543-8757         Sequence 8757, Ap
c 28     880      6.7    718  22  US-10-278-698-515          Sequence 515, App
c 29     880      6.7    718  22  US-10-278-698-1031         Sequence 1031, Ap
  30   746.5      5.6    520   9  US-09-920-300A-485        Sequence 485, App
  31   746.5      5.6    520  14  US-10-033-528-485          Sequence 485, App
  32   746.5      5.6    520  17  US-10-099-926-485          Sequence 485, App
  33   746.5      5.6    520  24  US-10-961-527-485          Sequence 485, App
c 34   730.5      5.5  29467  22  US-10-937-730A-1           Sequence 1, Appli
c 35   730.5      5.5  67167  22  US-10-937-730A-3           Sequence 3, Appli
  36     730      5.5 9025608  16  US-10-156-761-1            Sequence 1, Appli
  37     721      5.5  88400  22  US-10-844-716-1            Sequence 1, Appli
c 38   716.5      5.4  27705  10  US-09-942-025-12           Sequence 12, Appl
c 39   716.5      5.4  67311  10  US-09-942-025-1            Sequence 1, Appli
c 40     700      5.3  13416  16  US-10-205-032-11           Sequence 11, Appl
c 41     700      5.3  60196  16  US-10-205-032-1            Sequence 1, Appli
c 42   689.5      5.2  64492  19  US-10-378-083-1            Sequence 1, Appli
  43     689      5.2    427  24  US-10-779-543-12895        Sequence 12895, A
  44     685      5.2  30690  20  US-10-204-862A-1           Sequence 1, Appli
  45     685      5.2  30690  26  US-11-005-196-1            Sequence 1, Appli

                                 ALIGNMENTS

RESULT 1
US-10-414-692-14
; Sequence 14, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
;  APPLICANT: X-Ceptor Therapeutics, Inc.
;  TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
;  TITLE OF INVENTION:  profile
;  FILE REFERENCE: 8012-002-US
;  CURRENT APPLICATION NUMBER: US/10/414,692
;  CURRENT FILING DATE:  2003-04-14
;  PRIOR APPLICATION NUMBER: 60/372,650
;  PRIOR FILING DATE: 2002-04-15
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
;   LENGTH: 7554
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-414-692-14

Alignment Scores:
Pred. No.:              0               Length:        7554
Score:                  13215.00         Matches:       2517
Percent Similarity:     100.00%         Conservative:  0
Best Local Similarity:  100.00%         Mismatches:    0
Query Match:            100.00%         Indels:        0
DB:                     18              Gaps:          0

US-09-522-753-5 (1-2517) x US-10-414-692-14 (1-7554)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840
```

```
Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 900

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 960

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1020

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1080

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1140

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1200

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1260

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1320

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1380

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1440

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1500

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1560

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1620

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1680

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1740

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1800

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1860

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1920

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1980

Qy         661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2040

Qy         681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2100

Qy         701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2160

Qy         721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2220

Qy         741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2280

Qy         761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2340

Qy         781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2400

Qy         801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2460

Qy         821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2520

Qy         841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2580

Qy         861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2640

Qy         881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2700

Qy         901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2760

Qy         921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2820

Qy         941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2880

Qy         961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2940

Qy         981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3000

Qy        1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3060

Qy        1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3120

Qy        1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3180

Qy        1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3240

Qy        1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3300

Qy        1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3360

Qy        1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3420

Qy        1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3421 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3480

Qy        1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3481 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3540

Qy        1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3541 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3600

Qy        1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3660

Qy        1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3661 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3720

Qy        1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3780

Qy        1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3840

Qy        1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3900

Qy        1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3960

Qy        1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3961 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4020

Qy        1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4021 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4080

Qy        1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4081 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4140
```

```
Qy        1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4141 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4200

Qy        1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4201 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4260

Qy        1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4261 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4320

Qy        1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4321 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4380

Qy        1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4381 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4440

Qy        1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4441 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4500

Qy        1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4501 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4560

Qy        1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4561 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4620

Qy        1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4621 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4680

Qy        1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4681 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4740

Qy        1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4741 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4800

Qy        1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4801 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4860

Qy        1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4861 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4920

Qy        1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4921 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4980

Qy        1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4981 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5040

Qy        1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5041 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5100

Qy        1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5101 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5160

Qy        1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5161 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5220

Qy        1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5221 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5280

Qy        1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5281 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5340

Qy        1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5341 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5400

Qy        1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5401 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5460

Qy        1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5461 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5520

Qy        1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5521 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5580

Qy        1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5581 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5640

Qy        1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5641 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5700

Qy        1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5701 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5760

Qy        1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5761 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5820

Qy        1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5821 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5880

Qy        1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5881 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5940

Qy        1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5941 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6000

Qy        2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6001 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6060

Qy        2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6061 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6120

Qy        2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6121 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6180

Qy        2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6181 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6240

Qy        2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6241 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6300

Qy        2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6301 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6360

Qy        2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6361 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6420

Qy        2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6421 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6480

Qy        2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6481 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6540

Qy        2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6541 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6600

Qy        2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6601 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6660

Qy        2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6661 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6720

Qy        2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6721 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6780

Qy        2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6781 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6840

Qy        2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6841 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6900

Qy        2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6901 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6960

Qy        2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6961 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7020

Qy        2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7021 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7080

Qy        2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7081 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7140

Qy        2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7141 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7200

Qy        2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7201 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7260

Qy        2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7261 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7320

Qy        2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7321 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7380

Qy        2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        7381 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7440

Qy        2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7441 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7500

Qy        2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7501 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7551

RESULT 2
US-10-174-014-11
; Sequence 11, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT:  Susan M. Freier
;  APPLICANT:  Kenneth W. Dobie
;  TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
;  FILE REFERENCE: PTS-0012
;  CURRENT APPLICATION NUMBER: US/10/174,014
;  CURRENT FILING DATE:  2002-06-17
;  NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 11
;   LENGTH: 8561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2)...(7555)
US-10-174-014-11

Alignment Scores:
Pred. No.:              0             Length:        8561
Score:                  13215.00       Matches:       2517
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     18            Gaps:          0

US-09-522-753-5 (1-2517) x US-10-174-014-11 (1-8561)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 181

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481
```

```
Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 901

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 961

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1021

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1501

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy        781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2401

Qy        801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2461

Qy        821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2521

Qy        841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2522 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2581

Qy        861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2641

Qy        881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2701

Qy        901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2761

Qy        921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2821

Qy        941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2881

Qy        961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2941

Qy        981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2942 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3001

Qy       1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3061

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3121

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3181

Qy       1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3241

Qy       1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3301

Qy       1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3361

Qy       1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3421

Qy       1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3481

Qy       1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3541

Qy       1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3601

Qy       1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3661

Qy       1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3721

Qy       1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3722 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3781

Qy       1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3782 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3841

Qy       1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3842 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3901

Qy       1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3902 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3961

Qy       1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3962 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4021

Qy       1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4022 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4081

Qy       1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4082 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4141

Qy       1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4142 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4201

Qy       1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4261

Qy       1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4262 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4321

Qy       1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4322 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4381

Qy       1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4382 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4441

Qy       1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4442 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4501

Qy       1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4502 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4561

Qy       1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4562 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4621

Qy       1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4622 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4681

Qy       1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4682 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4741

Qy       1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4742 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4801

Qy       1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4802 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4861
```

```
Qy   1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4862 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4921

Qy   1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4922 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4981

Qy   1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4982 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5041

Qy   1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5042 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5101

Qy   1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5102 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5161

Qy   1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5162 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5221

Qy   1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5222 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5281

Qy   1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5282 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5341

Qy   1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5342 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5401

Qy   1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5402 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5461

Qy   1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5462 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5521

Qy   1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5522 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5581

Qy   1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5582 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5641

Qy   1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5642 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5701

Qy   1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5702 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5761

Qy   1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5762 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5821

Qy   1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5822 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5881

Qy   1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5882 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5941

Qy   1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5942 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6001

Qy   2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6002 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6061

Qy   2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6062 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6121

Qy   2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6122 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6181

Qy   2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6182 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6241

Qy   2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6242 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6301

Qy   2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6302 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6361

Qy   2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6362 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6421

Qy   2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6422 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6481

Qy   2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6482 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6541

Qy   2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6542 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6601

Qy   2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6602 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6661

Qy   2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6662 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6721

Qy   2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6722 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6781

Qy   2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6782 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6841

Qy   2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6842 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6901

Qy   2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6902 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6961

Qy   2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6962 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7021

Qy   2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7022 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7081

Qy        2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7082 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7141

Qy        2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7142 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7201

Qy        2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7202 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7261

Qy        2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7262 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7321

Qy        2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7322 CGCCGGACGCCGCTCACCAACCGCGTGTGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7381

Qy        2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7382 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7441

Qy        2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7442 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7501

Qy        2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7502 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7552

RESULT 3
US-11-052-344-1
; Sequence 1, Application US/11052344
; Publication No. US20050191674A1
; GENERAL INFORMATION:
;  APPLICANT: Coignet, Lionel
;  TITLE OF INVENTION: Method Of Prognosis Of Metastasis By Detection Of FRA12E Fragile
;  TITLE OF INVENTION:  Within The SMRT Gene/locus At Chromosome 12q24
;  FILE REFERENCE: 03551.0181
;  CURRENT APPLICATION NUMBER: US/11/052,344
;  CURRENT FILING DATE:  2005-02-07
;  PRIOR APPLICATION NUMBER: US 60/542,538
;  PRIOR FILING DATE: 2004-02-06
;  NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
;   LENGTH: 8561
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: SMRT gene which includes the open reading frame for SMRT protein.
;   OTHER INFORMATION: The open reading frame starts at nucleotide 2 and ends at nucleot
US-11-052-344-1

Alignment Scores:
Pred. No.:                 0              Length:          8561
Score:                     13215.00        Matches:         2517
Percent Similarity:        100.00%        Conservative:    0
Best Local Similarity:     100.00%        Mismatches:      0
Query Match:               100.00%        Indels:          0
DB:                        26             Gaps:            0

US-09-522-753-5 (1-2517) x US-11-052-344-1 (1-8561)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 181

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy         161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy         181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy         201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy         221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy         241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy         261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy         281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 901

Qy         301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         902 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 961

Qy         321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         962 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1021

Qy         341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy         361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141

Qy         381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1501

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2282 GACACAGGGCAGAATGGCCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGCCCACCCCCA 2341

Qy        781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2342 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2401

Qy        801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2402 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2461

Qy        821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2462 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2521

Qy        841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2522 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2581

Qy        861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2582 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2641

Qy        881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2642 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2701

Qy        901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2702 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2761

Qy        921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2762 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2821

Qy        941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2822 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2881

Qy        961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2941

Qy        981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2942 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3001

Qy       1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3002 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3061

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3062 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3121

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3122 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3181

Qy       1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3182 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3241

Qy       1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3242 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3301

Qy       1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3302 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3361
```

```
Qy       1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3362 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3421

Qy       1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3422 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3481

Qy       1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3482 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3541

Qy       1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3542 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3601

Qy       1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3602 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3661

Qy       1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3662 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3721

Qy       1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3722 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3781

Qy       1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3782 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3841

Qy       1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3842 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3901

Qy       1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3902 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3961

Qy       1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3962 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4021

Qy       1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4022 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4081

Qy       1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4082 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4141

Qy       1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4142 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4201

Qy       1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4261

Qy       1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4262 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4321

Qy       1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4322 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4381

Qy       1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4382 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4441

Qy       1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4442 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4501

Qy       1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4502 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4561

Qy       1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4562 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4621

Qy       1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4622 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4681

Qy       1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4682 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4741

Qy       1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4742 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4801

Qy       1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4802 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4861

Qy       1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4862 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4921

Qy       1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4922 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4981

Qy       1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4982 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5041

Qy       1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5042 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5101

Qy       1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5102 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5161

Qy       1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5162 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5221

Qy       1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5222 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5281

Qy       1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5282 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5341

Qy       1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5342 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5401

Qy       1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5402 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5461

Qy       1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5462 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5521

Qy       1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5522 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5581

Qy          1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5582 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5641

Qy          1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5642 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5701

Qy          1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5702 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5761

Qy          1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5762 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5821

Qy          1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5822 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5881

Qy          1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5882 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5941

Qy          1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5942 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6001

Qy          2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6002 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6061

Qy          2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6062 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6121

Qy          2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6122 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6181

Qy          2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6182 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6241

Qy          2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6242 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6301

Qy          2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6302 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6361

Qy          2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6362 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6421

Qy          2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6422 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6481

Qy          2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6482 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6541

Qy          2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6542 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6601

Qy          2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6602 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6661

Qy          2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6662 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6721

Qy          2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6722 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6781

Qy          2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6782 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6841

Qy          2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6842 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6901

Qy          2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6902 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6961

Qy          2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6962 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7021

Qy          2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7022 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7081

Qy          2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7082 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7141

Qy          2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7142 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7201

Qy          2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7202 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7261

Qy          2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7262 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7321

Qy          2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7322 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7381

Qy          2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7382 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7441

Qy          2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7442 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7501

Qy          2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7502 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7552

RESULT 4
US-10-087-192-653
; Sequence 653, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 653
;   LENGTH: 8667
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-087-192-653

Alignment Scores:
Pred. No.:              0            Length:          8667
Score:                  13145.00      Matches:         2509
Percent Similarity:     99.68%       Conservative:    1
Best Local Similarity:  99.64%       Mismatches:      7
Query Match:            99.47%       Indels:          2
DB:                     14           Gaps:            0

US-09-522-753-5 (1-2517) x US-10-087-192-653 (1-8667)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db         116 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 175

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         176 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 235

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         236 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 295

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         296 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 355

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         356 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 415

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         416 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 475

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         476 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 535

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         536 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 595

Qy         161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         596 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 655

Qy         181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         656 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 715

Qy         201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         716 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 775

Qy         221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         776 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 835

Qy         241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         836 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 895

Qy         261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         896 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 955

Qy         281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
               ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db         956 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1015

Qy         301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
               ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        1016 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1075

Qy         321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1076 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1135

Qy         341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1136 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1195

Qy         361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1196 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1255

Qy         381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1256 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1315

Qy         401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1316 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1375

Qy         421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1376 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1435

Qy         441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1436 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1495

Qy         461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1496 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1555

Qy         481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1556 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1615

Qy         501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1616 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1675

Qy         521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1676 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1735

Qy         541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1736 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1795

Qy         561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1796 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1855

Qy         581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1856 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1915

Qy         601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1916 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1975
```

```
Qy     621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1976 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2035

Qy     641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2036 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2095

Qy     661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2096 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2155

Qy     681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2156 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2215

Qy     701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2216 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2275

Qy     721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2276 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2335

Qy     741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2336 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2395

Qy     761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2396 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGCCACCCCCA 2455

Qy     781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
           |||||||||||||||||| |||||||||||||||    |||  |||||||||||||||||
Db    2456 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCGGCCTC 2514

Qy     800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2515 TGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCC 2574

Qy     820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2575 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2634

Qy     840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2635 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2694

Qy     860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2695 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2754

Qy     880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2755 GGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2814

Qy     900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2815 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTAC 2874

Qy     920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2875 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2934

Qy     940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2935 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2994

Qy     960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2995 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3054

Qy     980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3055 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3114

Qy    1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3115 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3174

Qy    1020 gGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGl 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3175 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCA 3234

Qy    1040 nLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProAr 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3235 GAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCG 3294

Qy    1060 gGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPr 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3295 TGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACC 3354

Qy    1080 oGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPr 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3355 TGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACC 3414

Qy    1100 oThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluAr 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3415 CACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAG 3474

Qy    1120 gGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHi 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3475 GCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCA 3534

Qy    1140 sAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLy 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3535 TGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAA 3594

Qy    1160 sLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPr 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3595 GCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCC 3654

Qy    1180 oProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLe 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3655 ACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCT 3714

Qy    1200 uGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAs 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3715 GGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGA 3774

Qy    1220 pSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLy 1240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3775 CAGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAA 3834

Qy    1240 sGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGl 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3835 GGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGA 3894

Qy    1260 uAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTy 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3895 GGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTA 3954

Qy    1280 rGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPr 1300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3955 TGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACC 4014

Qy    1300 oProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyAr 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4015 CCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAG 4074

Qy    1320 gAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHi 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4075 AGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACA 4134

Qy    1340 sSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePr 1360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4135 CAGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCC 4194

Qy        1360 oArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysAr 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4195 TCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCG 4254

Qy        1380 gGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGl 1400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4255 GGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCA 4314

Qy        1400 nAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGl 1420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4315 GGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGA 4374

Qy        1420 uAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPr 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4375 GGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCC 4434

Qy        1440 oLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAs 1460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4435 CCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGA 4494

Qy        1460 pThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPr 1480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4495 CACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCC 4554

Qy        1480 oGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGl 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4555 CGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGA 4614

Qy        1500 uArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGl 1520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4615 ACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGG 4674

Qy        1520 ySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPr 1540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4675 CTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCCGCGGCAGAGCCC 4734

Qy        1540 oLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVa 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4735 CCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGT 4794

Qy        1560 lThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSe 1580
               ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4795 GACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATC 4854

Qy        1580 rGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVa 1600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4855 CCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGT 4914

Qy        1600 lProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGl 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4915 GCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGG 4974

Qy        1620 yValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGl 1640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4975 CGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGG 5034

Qy        1640 yIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProTh 1660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5035 CATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCAC 5094

Qy        1660 rTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGl 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5095 CTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGA 5154

Qy        1680 uAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAl 1700
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db        5155 GAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGC 5214

Qy        1700 aThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLe 1720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5215 CACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCT 5274

Qy        1720 uAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLe 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5275 GGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCT 5334

Qy        1740 uProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5335 GCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTA 5394

Qy        1760 rLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGl 1780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5395 CCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGG 5454

Qy        1780 yGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspAr 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5455 AGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCG 5514

Qy        1800 gAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrTh 1820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5515 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGAC 5574

Qy        1820 rValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySe 1840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5575 GGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAG 5634

Qy        1840 rSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHi 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5635 CAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCA 5694

Qy        1860 sSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAs 1880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5695 CTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAA 5754

Qy        1880 nThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSe 1900
               ||||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db        5755 CACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGGTC 5814

Qy        1900 rThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPr 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5815 CACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCC 5874

Qy        1920 oLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLy 1940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5875 ACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAA 5934

Qy        1940 sGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLe 1960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5935 GGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCT 5994

Qy        1960 uAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGl 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5995 CGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGA 6054

Qy        1980 uProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLy 2000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6055 GCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAA 6114

Qy        2000 sAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAs 2020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6115 GAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGA 6174

Qy        2020 pProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6175 CCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTC 6234

Qy        2040 rLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSe 2060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6235 TCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAG 6294

Qy        2060 rSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSe 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6295 CTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAG 6354
```

```
Qy       2080 rHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAl 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6355 CCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGC 6414

Qy       2100 aAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLe 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6415 CGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCT 6474

Qy       2120 uGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSe 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6475 CCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAG 6534

Qy       2140 rGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPr 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6535 TGAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCC 6594

Qy       2160 oAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPr 2180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6595 CGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACC 6654

Qy       2180 oSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSe 2200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6655 CAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAG 6714

Qy       2200 rGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAs 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6715 CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGA 6774

Qy       2220 pGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAl 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6775 CGGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGC 6834

Qy       2240 aValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLy 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6835 TGTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAA 6894

Qy       2260 sSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSe 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6895 GTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTC 6954

Qy       2280 rAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAs 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6955 CGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAA 7014

Qy       2300 nGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleTh 2320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7015 TGAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCAC 7074

Qy       2320 rGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMe 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7075 CGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACAT 7134

Qy       2340 tGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSe 2360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7135 GGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTC 7194

Qy       2360 rProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAl 2380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7195 CCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGC 7254

Qy       2380 aMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGl 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7255 TATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGG 7314

Qy       2400 yGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGl 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7315 CGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGG 7374

Qy       2420 yLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAs 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7375 CCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAA 7434

Qy       2440 nArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerTh 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7435 CCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCAC 7494

Qy       2460 rProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPr 2480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7495 GCCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACC 7554

Qy       2480 oProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGl 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7555 CCCACCGGGCCTCCCCGCGGGCAGCGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGA 7614

Qy       2500 uGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7615 GGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7666

RESULT 5
US-09-819-104A-3
; Sequence 3, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, J. Don
;  TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: UMG-030
;  CURRENT APPLICATION NUMBER: US/09/819,104A
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/193,138
;  PRIOR FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 7521
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(7521)
US-09-819-104A-3

Alignment Scores:
Pred. No.:              0               Length:        7521
Score:                  12978.00         Matches:       2487
Percent Similarity:     98.57%          Conservative:  3
Best Local Similarity:  98.46%          Mismatches:    9
Query Match:            98.21%          Indels:        28
DB:                     10              Gaps:          3

US-09-522-753-5 (1-2517) x US-09-819-104A-3 (1-7521)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db          1 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360
```

```
Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 900

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        901 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1020

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1077

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1078 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1137

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1197

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1257

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1317

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1377

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1437

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1497

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1557

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1617

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1677

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1737

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1738 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1797

Qy        601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1798 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1857

Qy        621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1858 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1917

Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1918 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1977

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1978 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2037

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2038 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2097

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db       2098 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2157

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||                                                   |||
Db       2158 GCTGAA---------------------------------------------------GCC 2166

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2167 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2226

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2286

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db       2287 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2345

Qy        800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
              ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db       2346 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2405

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2406 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2465

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2466 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2525

Qy        860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2526 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2585

Qy        880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       2586 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2645

Qy        900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2646 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2705

Qy        920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2706 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2765

Qy        940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2766 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2825

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2826 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 2885

Qy        980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2886 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 2945

Qy       1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2946 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3005

Qy       1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
              ||||||||||||||||||||||||||||||||||||||||
Db       3006 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3065

Qy       1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3066 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3125

Qy       1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3126 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3185

Qy       1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3186 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3245

Qy       1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3246 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3305

Qy       1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3306 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3365

Qy       1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3366 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3425

Qy       1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3426 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3485

Qy       1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3486 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3545

Qy       1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3546 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3605

Qy       1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3606 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3665

Qy       1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3666 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3725

Qy       1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3726 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3785

Qy       1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3786 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 3845

Qy       1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3846 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 3905

Qy       1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3906 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 3965

Qy       1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3966 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4025

Qy       1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4026 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4085

Qy       1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4086 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4145

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4146 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4205

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4206 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4265

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4266 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4325

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4326 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4385

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4386 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4445

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4446 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4505

Qy       1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4506 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4565

Qy       1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4566 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4625

Qy       1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4626 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4685
```

```
Qy        1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4686 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4745

Qy        1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4746 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4805

Qy        1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4806 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 4865

Qy        1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4866 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 4925

Qy        1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4926 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 4985

Qy        1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4986 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5045

Qy        1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5046 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5105

Qy        1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5106 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5165

Qy        1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5166 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5225

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5226 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5285

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5286 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5345

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5346 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5405

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5406 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5465

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5466 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5525

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5526 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5585

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5586 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5645

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5646 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5705

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5706 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5765

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5766 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5825

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5826 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 5885

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5886 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 5945

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5946 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6005

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6006 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6065

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6066 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6125

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6126 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6185

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6186 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6245

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6246 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6305

Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6306 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6365

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6366 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6425

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6426 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6485

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6486 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6545

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6546 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6605

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6606 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6665

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6666 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6725

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6726 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6785

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6786 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 6845

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6846 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGACGGA 6905

Qy    2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6906 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 6965

Qy    2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    6966 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7025

Qy    2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7026 TAAATATGACCAGTGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7085

Qy    2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7086 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7145

Qy    2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7146 CACACTCACCTCGCCAGGTGGCGGCGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7205

Qy    2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7206 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7265

Qy    2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7266 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7325

Qy    2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7326 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7385

Qy    2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7386 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7445

Qy    2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    7446 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7505

Qy    2512 rLeuSerAspSerGlu 2517
           ||||||||||||||||
Db    7506 ACTCTCCGACAGCGAG 7521

RESULT 6
US-09-819-104A-1
; Sequence 1, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, J. Don
;  TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: UMG-030
;  CURRENT APPLICATION NUMBER: US/09/819,104A
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/193,138
;  PRIOR FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (157)..(7677)
US-09-819-104A-1

Alignment Scores:
Pred. No.:               0             Length:        8686
Score:                   12978.00       Matches:       2487
```

```
Percent Similarity:       98.57%          Conservative:   3
Best Local Similarity:  98.46%          Mismatches:     9
Query Match:              98.21%          Indels:         28
DB:                       10              Gaps:           3

US-09-522-753-5 (1-2517) x US-09-819-104A-1 (1-8686)

Qy       1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
           ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db     157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy      21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy      41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy      61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy      81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy     101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy     121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy     141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy     161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy     181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy     201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy     221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy     241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy     261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy     281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
           ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db     997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1056

Qy     301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db    1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy     321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy         341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
               |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy         361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
               ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy         381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy         401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy         421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy         441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy         461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy         481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy         501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy         521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy         541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy         561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy         581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy         601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy         621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy         641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy         661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy         681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253
```

```
Qy         701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
               ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy         721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
               ||||||                                                   |||
Db        2314 GCTGAA---------------------------------------------------GCC 2322

Qy         741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy         761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy         781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
               |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db        2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy         800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
               ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db        2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy         820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy         840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy         860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy         880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
               |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy         900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy         920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy         940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy         960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy         980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy        1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy        1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
               ||||||||||||||||||||||||||||||||||||||||
Db        3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy        1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281
```

```
Qy       1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy       1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy       1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy       1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy       1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy       1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy       1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy       1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy       1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy       1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy       1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy       1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy       1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy       1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy       1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy       1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy       1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432

              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy       1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy       1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy       1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy       1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy       1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy       1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy       1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy       1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy       1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy       1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy       1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy       1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy       1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy       1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy          1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy          1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy          1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy          1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy          1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy          1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy          1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy          1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy          1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy          1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy          1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy          2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy          2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy          2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy          2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy          2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy          2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy          2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy          2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy          2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy          2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy          2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy          2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy          2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy          2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy          2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy          2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy          2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy          2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy          2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy          2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy          2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy          2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy          2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy          2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy          2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661
```

```
Qy       2512 rLeuSerAspSerGlu 2517
              |||||||||||||||
Db       7662 ACTCTCCGACAGCGAG 7677

RESULT 7
US-10-174-014-4
; Sequence 4, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT:  Susan M. Freier
;  APPLICANT:  Kenneth W. Dobie
;  TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
;  FILE REFERENCE: PTS-0012
;  CURRENT APPLICATION NUMBER: US/10/174,014
;  CURRENT FILING DATE:  2002-06-17
;  NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 4
;   LENGTH: 8686
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (157)...(7680)
US-10-174-014-4

Alignment Scores:
Pred. No.:                0                Length:        8686
Score:                    12978.00          Matches:       2487
Percent Similarity:       98.57%           Conservative:  3
Best Local Similarity:    98.46%           Mismatches:    9
Query Match:              98.21%           Indels:        28
DB:                       18               Gaps:          3

US-09-522-753-5 (1-2517) x US-10-174-014-4 (1-8686)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1056

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy         561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy         581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy         601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy         621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy         641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy         661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy         681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy         701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
               ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy         721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
               ||||||                                                   |||
Db        2314 GCTGAA---------------------------------------------------GCC 2322

Qy         741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy         761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy         781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
               |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db        2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy         800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
               ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db        2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy         820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy         840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy         860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy         880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
               |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db        2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy         900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy         920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy         940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy         960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy         980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy        1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy        1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
               ||||||||||||||||||||||||||||||||||||||||
Db        3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy        1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy        1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy        1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy        1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy        1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy        1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy        1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy        1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy        1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy        1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy        1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy        1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941
```

```
Qy          1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy          1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4002 GGACGGCAGAAGCAGCTCAGGACCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy          1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy          1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy          1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy          1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy          1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy          1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy          1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy          1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy          1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy          1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy          1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy          1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy          1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy          1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy          1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy          1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy          1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy          1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy          1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy          1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy          1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy          1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy          1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy          1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy          1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy          1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy          1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy          1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy          1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy          1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy          1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy          1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy          1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy          1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy          1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy          2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy          2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy          2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy          2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy          2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy          2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy          2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy          2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy          2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy          2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy          2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy          2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy          2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy          2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy          2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy          2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy          2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy          2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy          2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy          2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7302 CACACTCACCTCGCCAGGTGGCGGCGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy          2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy          2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGAGGA 7481

Qy          2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy          2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGCCCCTCGC 7601

Qy          2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy          2512 rLeuSerAspSerGlu 2517
                 ||||||||||||||||
Db          7662 ACTCTCCGACAGCGAG 7677

RESULT 8
US-10-723-860-1739
; Sequence 1739, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1739
;   LENGTH: 8686
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-1739

Alignment Scores:
Pred. No.:               0              Length:          8686
Score:                   12978.00        Matches:         2487
Percent Similarity:      98.57%         Conservative:    3
Best Local Similarity:   98.46%         Mismatches:      9
Query Match:             98.21%         Indels:          28
DB:                      21             Gaps:            3

US-09-522-753-5 (1-2517) x US-10-723-860-1739 (1-8686)

Qy             1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                 ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db           157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy            21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy           41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy           61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy           81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy          101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy          121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy          141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy          161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy          181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy          201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1056

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db         1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db         1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||                                                   |||
Db         2314 GCTGAA---------------------------------------------------GCC 2322

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382
```

```
Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy        781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
              |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db       2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy        800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
              ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db       2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy        820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy        840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy        860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy        880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy        900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy        920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy        940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy        980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy       1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161

Qy       1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
              ||||||||||||||||||||||||||||||||||||||||
Db       3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy       1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy       1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy       1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy       1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy       1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy       1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy       1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy       1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy       1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy       1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy       1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy       1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy       1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy       1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy       1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy       1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy       1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy       1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy       1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy       1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy       1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy       1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy       1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy       1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy       1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy       1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy       1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy       1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy       1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy       1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy       1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy       1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy       1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy       1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy       1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681
```

```
Qy       1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy       1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy       1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy       1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy       1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy       1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy       1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy       1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy       2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy       2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy       2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy       2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy       2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461

Qy       2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy       2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy       2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy       2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy       2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761
```

```
Qy       2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy       2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy       2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy       2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy       2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy       2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy       2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy       2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy       2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy       2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy       2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy       2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy       2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541

Qy       2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy       2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy       2512 rLeuSerAspSerGlu 2517
              ||||||||||||||||
Db       7662 ACTCTCCGACAGCGAG 7677

RESULT 9
US-10-723-860-6114/c
; Sequence 6114, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6114
;   LENGTH: 9079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-723-860-6114

Alignment Scores:
Pred. No.:              0               Length:         9079
Score:                  12845.50         Matches:        2473
Percent Similarity:     98.10%          Conservative:   6
Best Local Similarity:  97.86%          Mismatches:     20
Query Match:            97.20%          Indels:         31
DB:                     21              Gaps:           3

US-09-522-753-5 (1-2517) x US-10-723-860-6114 (1-9079)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       8552 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 8493

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8492 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 8433

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8432 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 8373

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8372 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 8313

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8312 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 8253

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8252 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 8193

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8192 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGGCC- 8134

Qy        141 LeuThrGlyLysLeuGluProValSer-ProProSerProProHisThrAspProGluLe 160
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       8133 CTGACGGGCAAGCTGGAACCGGTGTCTCCCACCCAGCCCCCCGCACACTGACCCTGAGCT 8074

Qy        160 uGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAs 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8073 GGAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGA 8014

Qy        180 pArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLe 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       8013 CCGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCT 7954

Qy        200 uGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGl 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7953 GGAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAACCCGTGTCACCGCCGCCCATCGA 7894

Qy        220 uSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAl 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7893 GTCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGC 7834

Qy        240 aAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPr 260
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7833 TGCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCC 7774

Qy          260 oSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLe 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7773 CTCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCT 7714

Qy          280 uIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGl 300
                |||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db         7713 AATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCA 7654

Qy          300 nArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPr 320
                |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db         7653 GCGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCC 7594

Qy          320 oArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIl 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7593 CCGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGAT 7534

Qy          340 eArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGl 360
                ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db         7533 CCGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGG 7477

Qy          360 yLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSe 380
                |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db         7476 GCTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTC 7417

Qy          380 rGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTy 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7416 AGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTA 7357

Qy          400 rAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMe 420
                ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db         7356 CGACGCTGACCAGCAGCGCATCACGTTCATCAACATGAACGGGCTTATGGCCGACCCCAT 7297

Qy          420 tLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPh 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7296 GAAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTT 7237

Qy          440 eArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluAr 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7236 CCGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAG 7177

Qy          460 gLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLy 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7176 GAAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAA 7117

Qy          480 sSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGl 500
                |||||||||||||||||||||||||||||||||||||:::   |||||||||   :::
Db         7116 GAGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAACC--CAGCAGCAAACAAAAGG 7059

Qy          500 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLy 520
                 :::|||                  |||||||||||||||||||||||||||||||||||
Db         7058 CCGGCAGGCAGCAGCCACCAGCAGCCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAA 6999

Qy          520 sAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAs 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6998 AGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAA 6939

Qy          540 nAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGl 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6938 CGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGA 6879

Qy          560 uLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGl 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6878 GAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGG 6819

Qy          580 yArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGl 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6818 CCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCA 6759

Qy          600 nSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMe 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         6758 GAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAAT 6699

Qy          620 tGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMe 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6698 GGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGAT 6639

Qy          640 tValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGl 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6638 GGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCA 6579

Qy          660 nAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaAr 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6578 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCG 6519

Qy          680 gArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVa 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6518 GAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGT 6459

Qy          700 lGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGl 720
                |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db         6458 GGAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGA 6399

Qy          720 uAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAl 740
                |||||||                                                   ||
Db         6398 GGCTGAA---------------------------------------------------GC 6390

Qy          740 aThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLy 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6389 CACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAA 6330

Qy          760 sAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPr 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6329 GGACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCC 6270

Qy          780 oGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaS 800
                ||||||||||||||||||| |||||||||||||||   |||   |||   ||||||||||
Db         6269 AGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCT 6211

Qy          800 erGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProP 820
                |||||||||||   |||||||||||||||||||||||||||   ||||||:::|||||||
Db         6210 CTGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTC 6151

Qy          820 roValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluG 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6150 CTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGG 6091

Qy          840 lyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluG 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6090 GGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGG 6031

Qy          860 luProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspA 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6030 AGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACG 5971

Qy          880 laGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlyS 900
                ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db         5970 CGGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGA 5911

Qy          900 erGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaT 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5910 GCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCA 5851

Qy          920 hrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerP 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5850 CCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCC 5791

Qy          940 roArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysP 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5790 CAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGC 5731

Qy          960 roLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrL 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5730 CACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCA 5671
```

```
Qy        980 ysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProP 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5670 AAGTCCATGAGCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCAC 5611

Qy       1000 roProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5610 CGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCC 5551

Qy       1020 rgGlyLysSerArgSerProAlaProProAlaAspLysGlu------------------- 1033
              |||||||||||||||||||||||||||||||||||||||||
Db       5550 GGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCT 5491

Qy       1034 -----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerG 1052
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5490 TCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCG 5431

Qy       1052 lyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspP 1072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5430 GCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACC 5371

Qy       1072 roSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrA 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5370 CCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTG 5311

Qy       1092 laArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerA 1112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5310 CCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTG 5251

Qy       1112 laLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValG 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5250 CCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCC 5191

Qy       1132 lnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyL 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5190 AGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGC 5131

Qy       1152 euProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnL 1172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5130 TGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGC 5071

Qy       1172 euSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluA 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5070 TGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGG 5011

Qy       1192 laSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyI 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5010 CGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCA 4951

Qy       1212 leProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisG 1232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4950 TTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACG 4891

Qy       1232 lyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerP 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4890 GCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCC 4831

Qy       1252 roSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluG 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4830 CGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAG 4771

Qy       1272 lyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysG 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4770 GCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGG 4711

Qy       1292 luAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrA 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4710 AGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATG 4651

Qy       1312 spMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetG 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4650 ACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGG 4591
```

```
Qy       1332 lyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleA 1352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4590 GCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCC 4531

Qy       1352 rgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuA 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4530 GCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGC 4471

Qy       1372 rgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgA 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4470 GTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGG 4411

Qy       1392 spLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisG 1412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4410 ACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATG 4351

Qy       1412 luGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluG 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4350 AGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGG 4291

Qy       1432 luLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleT 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4290 AGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCA 4231

Qy       1452 hrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisA 1472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4230 CGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACG 4171

Qy       1472 spValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspV 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4170 ACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATG 4111

Qy       1492 alMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgP 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4110 TGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGC 4051

Qy       1512 roGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProG 1532
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4050 CAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTG 3991

Qy       1532 luLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaG 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3990 AGCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCG 3931

Qy       1552 lyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluG 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3930 GCCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGG 3871

Qy       1572 lySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluI 1592
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3870 GCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGA 3811

Qy       1592 leAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrG 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3810 TCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATG 3751

Qy       1612 luHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaP 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3750 AGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCT 3691

Qy       1632 heAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuP 1652
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3690 TCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGC 3631

Qy       1652 roArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgG 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3630 CCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCG 3571

Qy       1672 lyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrS 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3570 GCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCT 3511

Qy       1692 erGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgG 1712
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CGCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGG 3451

Qy       1712 lyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleI 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3450 GCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCA 3391

Qy       1732 leAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProA 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3390 TCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAG 3331

Qy       1752 laThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgH 1772
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3330 CCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCC 3271

Qy       1772 isSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrS 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3270 ACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGT 3211

Qy       1792 erSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluL 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3210 CCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAA 3151

Qy       1812 ysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrG 1832
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3150 AGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAG 3091

Qy       1832 luGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProA 1852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3090 AGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCG 3031

Qy       1852 laSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuG 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3030 CCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCC 2971

Qy       1872 lnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluP 1892
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2970 AGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGC 2911

Qy       1892 roSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaT 1912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2910 CCAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCA 2851

Qy       1912 hrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrL 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2850 CATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCC 2791

Qy       1932 euMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProA 1952
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2790 TCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCC 2731

Qy       1952 rgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProA 1972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2730 GAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCG 2671

Qy       1972 laSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisA 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2670 CCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACG 2611

Qy       1992 laThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProP 2012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2610 CCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGC 2551

Qy       2012 roAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheS 2032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2550 CGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTT 2491

Qy       2032 erIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluG 2052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2490 CCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAG 2431

Qy       2052 lyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProL 2072
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2430 GGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCA 2371

Qy       2072 ysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProG 2092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2370 AGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAG 2311

Qy       2092 lyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluS 2112
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2310 GCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGA 2251

Qy       2112 erGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgV 2132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2250 GCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGG 2191

Qy       2132 alValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisP 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2190 TGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACC 2131

Qy       2152 roGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysP 2172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2130 CACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCC 2071

Qy       2172 roValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyA 2192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2070 CCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTG 2011

Qy       2192 laProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysT 2212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2010 CCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGA 1951

Qy       2212 hrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetT 2232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1950 CGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGA 1891

Qy       2232 hrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnT 2252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1890 CGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGA 1831

Qy       2252 hrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPheP 2272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1830 CGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCT 1771

Qy       2272 heSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnL 2292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1770 TCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACA 1711

Qy       2292 ysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrG 2312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1710 AGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGG 1651

Qy       2312 luIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaV 2332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1650 AGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGG 1591

Qy       2332 alGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetG 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1590 TGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGG 1531

Qy       2352 lyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuA 2372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1530 GTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGA 1471

Qy       2372 snAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspH 2392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1470 ATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACC 1411

Qy       2392 isThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerA 2412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1410 ACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCC 1351

Qy       2412 rgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerS 2432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 GAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCT 1291
```

```
Qy       2432 erValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluA 2452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1290 CAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGG 1231

Qy       2452 spArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuG 2472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1230 ACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGC 1171

Qy       2472 lnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuA 2492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1170 AGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCG 1111

Qy       2492 laGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluT 2512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1110 CTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGA 1051

Qy       2512 hrLeuSerAspSerGlu 2517
              |||||||||||||||||
Db       1050 CACTCTCCGACAGCGAG 1034

RESULT 10
US-09-819-104A-6
; Sequence 6, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, J. Don
;  TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: UMG-030
;  CURRENT APPLICATION NUMBER: US/09/819,104A
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/193,138
;  PRIOR FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 7386
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(7386)
US-09-819-104A-6

Alignment Scores:
Pred. No.:              0              Length:        7386
Score:                  10832.50        Matches:       2149
Percent Similarity:     87.94%         Conservative:  90
Best Local Similarity:  84.41%         Mismatches:    197
Query Match:            81.97%         Indels:        117
DB:                     10             Gaps:          33

US-09-522-753-5 (1-2517) x US-09-819-104A-6 (1-7386)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db          1 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 60

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db         61 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 120

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db        121 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 180

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        181 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGGTCTCAG 240

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||   ::::::||||||||||||||||||   ::::::
Db        241 GAGCTCCACCTGCGCCCTGAGTCCCGCACGTTCCTGCCTGAGCTGGGCAAGCCCGACATA 300

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||   ||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        301 GAATTCACCGAGAGCAAGCGCCCCCGCCTGGAGCTACTACCCGATACCCTGCTGCGCCCA 360

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        361 TCACCCCTGCTGGCCACTGGGCAGCCGAGTGGGTCTGAAGACCTTACCAAGGACCGTAGC 420

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        421 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 480

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsn---MetAspArgVal 179
              ||||||   |||   ||||||||||||||||||||||||||||||   :::|||||||||
Db        481 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACAGATTGGACCGCGTG 540

Qy        180 AspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGln 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GACCGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAG 600

Qy        200 LeuGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIle 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTGGAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATA 660

Qy        220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAATCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGAAGAAAGCCGAA 720

Qy        240 AlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGln 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GCCGCACACCGGATCCTAGAAGGCCTGGGGCCCCAGGTGGAGCTGCCTCTGTACAACCAG 780

Qy        260 ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CCGTCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAG 840

Qy        280 LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys 299
              |||||||||||||||||||||||||||||||||||||||||||||:::|||:::||||||
Db        841 CTGATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGAACAGCGCTTCTGC 900

Qy        300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsn 319
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        901 CAGCGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAAT 960

Qy        320 ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CCGCGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAG 1020

Qy        340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ATCCGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGT 1080

Qy        360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GGGCTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTG 1140

Qy        380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       1141 TCTGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCC--GCCATGTTG 1198

Qy        400 TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro 419
              ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       1199 TACGACGC-GACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCC 1257

Qy        420 MetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThr 439
              |||||||||||||||||||||||||||   |||||||||||||||||||||::::::|||
Db       1258 ATGAAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACC 1317

Qy        440 PheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 TTCCGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAG 1377
```

```
Qy        460 ArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyr 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1378 AGAAAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTAC 1437

Qy        480 LysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGln 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1438 AAGAGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAG 1497

Qy        500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
              ||||||||||||||||||||||||            |||   ||||||||||||||||||
Db       1498 CAACAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAG 1545

Qy        520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
              |||:::||||||||||||||||||||||||:::|||||||||||||||   :::   |||
Db       1546 AAGGAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAG 1605

Qy        540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
              |||:::||||||:::|||   |||||||||||||||||||||||||||||||||||||||
Db       1606 AACGAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACGAT 1665

Qy        560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1666 GAGAAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAA 1725

Qy        580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
              ||||||||||||||||||||||||||||||||||||   ||||||      |||||||||
Db       1726 GGCCGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAG 1785

Qy        600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
              ||||||:::||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1786 CAAAGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAG 1845

Qy        620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1846 ATGGAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGC 1905

Qy        640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 ATGGTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 1965

Qy        660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1966 CAGAACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCT 2025

Qy        680 ArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProVal 699
              ||||||||||||||||||   |||||||||||||||||||||   ||||||||||||   
Db       2026 CGGAGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCC 2085

Qy        700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
                 |||||||||||||||||||||||||||   |||   ||||||||||||:::   |||
Db       2086 GCTGAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAG 2145

Qy        720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGly 738
              ||||||||||||      ||||||||||||||||||||||||   |||||||||||||||
Db       2146 GAGGCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGC 2205

Qy        739 ProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAla 758
              ||||||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db       2206 CCAGCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCC 2265

Qy        759 AlaLysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAsp 776
                 |||||||||         ||||||||||||      |||      |||   |||   
Db       2266 ACGAAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCC 2316

Qy        777 GlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluPro 796
                    ||||||   |||      ||||||         :::   ||||||   ||||||
Db       2317 ACCCAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGCAGCCCCTGCTGAGCCC 2370

Qy        797 ThrProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSer 816
              :::|||      :::|||:::|||   |||:::|||   |||:::   |||||||||:::
Db       2371 TCCCCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTCC---CCATCACCTGCC 2427

Qy        817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
              |||||||||      |||   |||:::|||:::|||         |||   ||||||   
Db       2428 GCACCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAG 2487

Qy        837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
                 |||:::   :::|||||||||      |||   |||         :::|||   |||
Db       2488 ACAGAAGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGGA 2538

Qy        857 LysAlaGluGluPro---------------------------ValLysSerGluCysThr 867
              |||   |||||||||                           |||||||||:::      
Db       2539 AAGCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTAAAGAGTGACCACAAG 2598

Qy        868 GluGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThr 886
              ||||||   ||||||   |||      |||   |||      ||||||   |||      
Db       2599 GAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTG 2658

Qy        887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThr--- 905
              :::|||      ||||||   |||      |||   ||||||      |||   |||   
Db       2659 TCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAGGCAGCTGTGACCAAG 2712

Qy        906 AlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGlu 925
                    ||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       2713 GGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACCTGCAGTGCCGATGAG 2772

Qy        926 ValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeu 945
              |||||||||   |||||||||||||||   ||||||||||||||||||||||||||||||
Db       2773 GTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTC 2832

Qy        946 ThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGln 965
              ||||||   |||||||||||||||:::   ||||||||||||||||||||||||||||||
Db       2833 ACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAG 2892

Qy        966 LeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProPro 985
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       2893 CTGAAGCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAGGTCCATGAGCCCCCC 2949

Qy        986 ArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeu 1005
              |||||||||      |||   ||||||   ||||||   |||||||||   |||:::|||
Db       2950 CGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTA 3009

Qy       1006 GlnProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSer 1025
              |||||||||   |||      ||||||   |||   ||||||||||||||||||||||||
Db       3010 CAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGC 3069

Qy       1026 ProAlaProProAlaAspLysGlu------------------------AlaPheAlaAla 1037
              |||   |||||||||:::||||||                        ||||||      
Db       3070 CCAGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACT 3129

Qy       1038 GluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProVal 1057
              |||      |||||||||   :::||||||   |||:::|||         ||||||   
Db       3130 GAGGGCCA-AAGCTACCGACTGAGCCCCCACGCTGGTCATCGGCT---GCCTTCCCAT-- 3183

Qy       1058 ProProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyr 1077
              |||||||||||||||||||||   |||      |||   ||||||      |||||||||
Db       3184 CCTCCACGGGAGGTGATCAAGACTTCCACACGCGCT---GACCCT---CTCTTCTCCTAC 3237

Qy       1078 AlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuPro 1097
                 ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       3238 ACACCCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCA 3297

Qy       1098 ArgProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerVal 1117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||   |||
Db       3298 CGTCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTA 3354

Qy       1118 LeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyr 1137
              ||||||||||||:::|||||||||||||||||||||||||||||||||   ||||||:::
Db       3355 CTTGAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCAC 3414

Qy       1138 SerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAsp 1157
              |||||||||||||||   |||:::||||||:::||||||   |||||||||   :::|||
Db       3415 TCAGAGCATGCCAAG---CCCATGGGCCCTCTCACCATGGAGCTGCCCCTTGCCGTGGAC 3471

Qy       1158 ProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGln 1177
```

```
              ||||||||||||
Db       3472 CCTAAGAAGCTG------------------------------------------------ 3483

Qy       1178 AlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGly 1197
                                                                       |||
Db       3484 ---------------------------------------------------------GGG 3486

Qy       1198 ThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVal 1217
              ||||||   ||||||         ||||||||||||||||||   ||||||||||||
Db       3487 ACAGCACT-GGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCT-CCCAGTACCCGGGCT 3544

Qy       1218 ProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVal 1237
                                :::|||||||||||||||||||||   |||||||||||||||
Db       3545 GCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGC-ACGCCCGCAGACGTC 3597

Qy       1238 LeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArg 1257
              ||||||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db       3598 CTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGG 3657

Qy       1258 GlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisVal 1277
                 |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       3658 GCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTC 3717

Qy       1278 LeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1297
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       3718 CTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAGC 3777

Qy       1298 SerGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArg 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3778 TCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCGT 3837

Qy       1318 ValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPro 1337
              |||||||||   ::::::||||||||||||||||||||||||||||||||||||   |||
Db       3838 GTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CCT 3894

Qy       1338 GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGln 1357
              |||:::|||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       3895 GAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAA 3951

Qy       1358 GlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3952 GGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGGAGGCCAAGCTC 4011

Qy       1378 LeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyr 1397
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||   |||
Db       4012 TTGAAGCGAGAAGGACACCACCTCCCCCACCACCACCTCGGGACCTGACTGAGACCTAC 4071

Qy       1398 LysThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeu 1414
              |||   :::            ||||||||||||||||||||||||   |||||||||:::
Db       4072 AAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGTG 4131

Qy       1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4132 GTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCGC 4191

Qy       1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4192 CGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGGC 4251

Qy       1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
              ||||||||||||||||||:::||||||   :::||||||:::||||||||||||||||||
Db       4252 ACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCGC 4311

Qy       1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
              |||:::||||||||||||||||||   ||||||   :::||||||||||||:::||||||
Db       4312 TCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCT 4371

Qy       1495 AspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThr 1514
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       4372 GACGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACC 4431

Qy       1515 AlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGly 1534
              :::|||   :::||||||||||||   |||||||||||||||:::|||||||||||||||
Db       4432 AGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGC 4491

Qy       1535 LysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeu 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       4492 AAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTG 4551

Qy       1555 ProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeu 1574
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       4552 CCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCTC 4611

Qy       1575 SerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLys 1594
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4612 CTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAAG 4671

Qy       1595 SerProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeu 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4672 TCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTTG 4731

Qy       1615 LeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPro 1634
              ||||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||
Db       4732 CTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCCC 4791

Qy       1635 ThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArg 1653
              |||||||||||||||||||||||||||:::   |||||||||||||||||||||||||||
Db       4792 ACCTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCGG 4851

Qy       1654 HisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyr 1673
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       4852 CACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTAC 4911

Qy       1674 ProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGln 1693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4912 CCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAG 4971

Qy       1694 GlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeu 1713
              |||||||||||||||   |||:::||||||||||||||||||||||||||||||||||||
Db       4972 CAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTG 5031

Qy       1714 SerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAsp 1733
              |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db       5032 TCACCGCGAGAGTCCTCGCTGGCCCTCAATTATTCCGCTGGCCCAAGAGGCATTATCGAC 5091

Qy       1734 LeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThr 1753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5092 CTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACC 5151

Qy       1754 AlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSer 1773
              |||:::||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db       5152 GCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGT 5211

Qy       1774 SerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSer 1793
              ||||||||||||||||||||||||||||||||||||   |||||||||   |||||||||
Db       5212 AGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCA 5271

Qy       1794 SerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
              |||||||||||||||:::|||:::||||||||||||                  ||||||
Db       5272 TCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTCC 5313

Qy       1814 IleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGln 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5314 ATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAG 5373

Qy       1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSer 1853
              |||||||||:::                        |||   ||||||||||||||||||
Db       5374 AGCAGCGGGGCT------------------------GGGGGCAGCAGCCGCCCCGCCTCC 5409

Qy       1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGln 1873
              |||:::      ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5410 CACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCAG 5463

Qy       1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
              ||||||||||||||||||||||||   |||||||||::::::|||:::|||||||||
Db       5464 AGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGC 5523
```

```
Qy       1894 LysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPhe 1913
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5524 ACGCCCACGGTCCTGAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGCCACATTC 5583

Qy       1914 ProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMet 1933
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db       5584 CCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTACCCTCATG 5643

Qy       1934 GluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAla 1953
              ||||||||||||||||||||||||      |||||||||||||||||||||||||||   
Db       5644 GAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGCCCCGTGTG 5703

Qy       1954 AspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSer 1973
              |||   |||||||||||||||   |||||||||||||||         ||||||||||||
Db       5704 GACGGTGGCCATGCCTTCCTCACCAAACCCCCGGCCCGG---------GAGCCCGCCTCC 5754

Qy       1974 SerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThr 1993
              ||||||||||||   ||||||||||||   |||   ||||||   |||   |||      
Db       5755 TCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCACACAGCC 5814

Qy       1994 IleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAla 2013
              |||||||||||||||||||||:::|||||||||||||||||||||||||||||||||   
Db       5815 ATCGCCCGCACCCCAGCAAAGAGCCTTGCACCCCACCATGCCAGTCCGGACCCGCCGGGG 5874

Qy       2014 ProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIle 2033
                 |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       5875 ---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTTTTCCATC 5931

Qy       2034 GlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSerProGluGly 2052
              |||||||||||||||||||||||||||||||||   |||:::   |||||||||:::|||
Db       5932 CAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAGCCCCGATGGG 5991

Qy       2053 ValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLys 2072
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db       5992 GTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGGGCTCTCCAAA 6051

Qy       2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
                 ||||||||||||:::|||||||||||||||||||||||||||   ||||||||||||
Db       6052 CCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAAGCAGCCAGGC 6111

Qy       2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSer 2112
              |||:::||||||      ||||||||||||||||||||||||||||||||||||||||||
Db       6112 CCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACTGCCCGAGAGC 6171

Qy       2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVal 2132
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       6172 CAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCAGAGGGTG 6231

Qy       2133 ValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPro 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6232 GTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACCCGGCACCACCCG 6291

Qy       2153 GlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPro 2172
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       6292 CAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGCCAGCTGCCCT 6351

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAla 2192
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db       6352 GTGCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGACCATGGCACC 6411

Qy       2193 ProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThr 2212
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       6412 CCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACCCAGCAAAACA 6471

Qy       2213 SerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThr 2232
              ||||||||||||      ||||||   |||||||||||||||||||||||||||||||||
Db       6472 TCGGTCCTGGGCAGCAGTGAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGGCATGACT 6531

Qy       2233 GluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThr 2252
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||   
Db       6532 GAGCCAGGACATGCTCGGAGCGCTGTGTACCCACTGCTGTATCGAGACGGGGAACAGGGC 6591

Qy       2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePhe 2272
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6592 GAGCCC---AGGATGGGCTCTAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTC 6648

Qy       2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLys 2292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6649 AGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGATCAACAAG 6708

Qy       2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGlu 2312
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db       6709 AAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCCTGGGACGGAA 6768

Qy       2313 IlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
              |||||||||||||||||||||||||||   ||||||||||||   |||||||||||||||
Db       6769 ATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAGCCAGGCGGTG 6828

Qy       2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6829 CAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGCACTCATGGGT 6888

Qy       2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
              |||||||||||||||||||||   |||||||||   ||||||||||||||||||||||||
Db       6889 AAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAACCCTCTGAAT 6948

Qy       2373 AlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
              ||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||||
Db       6949 GCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGGACGGAGTGAC 7008

Qy       2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7009 CACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTCTGGCAGACCTAGCAGC 7068

Qy       2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7069 CGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCCTTCTGTCTCC 7128

Qy       2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7129 TCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCGTGTGTGGGAG 7188

Qy       2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7189 GACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGATTATGAGGCTA 7248

Qy       2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       7249 CAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAGCGGGCCCCTA 7308

Qy       2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7309 GCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTCACAGTATGAG 7368

Qy       2512 ThrLeuSerAspSerGlu 2517
              ||||||||||||||||||
Db       7369 ACACTCTCGGACAGCGAG 7386

RESULT 11
US-09-819-104A-4
; Sequence 4, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, J. Don
;  TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: UMG-030
;  CURRENT APPLICATION NUMBER: US/09/819,104A
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/193,138
;  PRIOR FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 8544
;    TYPE: DNA
;    ORGANISM: Mus musculus
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (160)..(7545)
US-09-819-104A-4

Alignment Scores:
Pred. No.:                 0              Length:          8544
Score:                     10832.50        Matches:         2149
Percent Similarity:        87.94%         Conservative:    90
Best Local Similarity:     84.41%         Mismatches:      197
Query Match:               81.97%         Indels:          117
DB:                        10             Gaps:            33

US-09-522-753-5 (1-2517) x US-09-819-104A-4 (1-8544)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db        160 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 219

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db        220 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 279

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db        280 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 339

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        340 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGGTCTCAG 399

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||   ::::::||||||||||||||||||   ::::::
Db        400 GAGCTCCACCTGCGCCCTGAGTCCCGCACGTTCCTGCCTGAGCTGGGCAAGCCCGACATA 459

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||   ||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        460 GAATTCACCGAGAGCAAGCGCCCCCGCCTGGAGCTACTACCCGATACCCTGCTGCGCCCA 519

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        520 TCACCCCTGCTGGCCACTGGGCAGCCGAGTGGGTCTGAAGACCTTACCAAGGACCGTAGC 579

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        580 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 639

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsn---MetAspArgVal 179
              ||||||   |||   ||||||||||||||||||||||||||||||   :::|||||||||
Db        640 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACAGATTGGACCGCGTG 699

Qy        180 AspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGln 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GACCGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAG 759

Qy        200 LeuGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIle 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TTGGAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATA 819

Qy        220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 GAATCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGAAGAAAGCCGAA 879

Qy        240 AlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGln 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 GCCGCACACCGGATCCTAGAAGGCCTGGGGCCCCAGGTGGAGCTGCCTCTGTACAACCAG 939

Qy        260 ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CCGTCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAG 999

Qy        280 LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys 299
              |||||||||||||||||||||||||||||||||||||||||||||:::|||:::||||||
Db       1000 CTGATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGGAACAGCGCTTCTGC 1059

Qy        300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsn 319
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db       1060 CAGCGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAAT 1119

Qy        320 ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 CCGCGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAG 1179

Qy        340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 ATCCGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGT 1239

Qy        360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GGGCTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTG 1299

Qy        380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       1300 TCTGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCC--GCCATGTTG 1357

Qy        400 TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro 419
              ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       1358 TACGACGC-GACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCC 1416

Qy        420 MetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThr 439
              |||||||||||||||||||||||||||   |||||||||||||||||||||::::::|||
Db       1417 ATGAAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACC 1476

Qy        440 PheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 TTCCGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAG 1536

Qy        460 ArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyr 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 AGAAAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTAC 1596

Qy        480 LysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGln 499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1597 AAGAGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAG 1656

Qy        500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
              ||||||||||||||||||||||||            |||   ||||||||||||||||||
Db       1657 CAACAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAG 1704

Qy        520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
              |||:::||||||||||||||||||||||||:::|||||||||||||||   :::   |||
Db       1705 AAGGAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAG 1764

Qy        540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
              |||:::||||||:::|||   |||||||||||||||||||||||||||||||||||||||
Db       1765 AACGAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACGAT 1824

Qy        560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1825 GAGAAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAA 1884

Qy        580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
              ||||||||||||||||||||||||||||||||||||   ||||||      |||||||||
Db       1885 GGCCGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAG 1944

Qy        600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
              ||||||:::||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1945 CAAAGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAG 2004

Qy        620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2005 ATGGAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGC 2064

Qy        640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2065 ATGGTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2124

Qy        660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2125 CAGAACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCT 2184

Qy        680 ArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProVal 699
              ||||||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db       2185 CGGAGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCC 2244

Qy        700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
                 |||||||||||||||||||||||||||   |||   ||||||||||||:::   |||
Db       2245 GCTGAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAG 2304

Qy        720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGly 738
              ||||||||||||      ||||||||||||||||||||||||   |||||||||||||||
Db       2305 GAGGCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGC 2364

Qy        739 ProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAla 758
              ||||||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db       2365 CCAGCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCC 2424

Qy        759 AlaLysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAsp 776
                 |||||||||         ||||||||||||      |||      |||   |||
Db       2425 ACGAAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCC 2475

Qy        777 GlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluPro 796
                    ||||||   |||      ||||||         :::   ||||||   ||||||
Db       2476 ACCCAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGCAGCCCCTGCTGAGCCC 2529

Qy        797 ThrProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSer 816
              :::|||      :::|||:::|||   |||:::|||   |||:::   |||||||||:::
Db       2530 TCCCCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTCC---CCATCACCTGCC 2586

Qy        817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
              |||||||||      |||   |||:::|||:::|||         |||   ||||||
Db       2587 GCACCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAG 2646

Qy        837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
                 |||:::   :::|||||||||      |||   |||         :::|||   |||
Db       2647 ACAGAAGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGGA 2697

Qy        857 LysAlaGluGluPro---------------------------ValLysSerGluCysThr 867
              |||   |||||||||                           |||||||||:::
Db       2698 AAGCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTAAAGAGTGACCACAAG 2757

Qy        868 GluGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThr 886
              ||||||   ||||||   |||      |||   |||      ||||||   |||
Db       2758 GAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTG 2817

Qy        887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThr--- 905
              :::|||      ||||||   |||      |||   ||||||      |||   |||
Db       2818 TCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAGGCAGCTGTGACCAAG 2871

Qy        906 AlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGlu 925
                    ||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       2872 GGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACCTGCAGTGCCGATGAG 2931

Qy        926 ValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeu 945
              |||||||||   |||||||||||||||   ||||||||||||||||||||||||||||||
Db       2932 GTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTC 2991

Qy        946 ThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGln 965
              ||||||   |||||||||||||||:::   ||||||||||||||||||||||||||||||
Db       2992 ACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAG 3051

Qy        966 LeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProPro 985
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       3052 CTGAAGCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAGGTCCATGAGCCCCCC 3108

Qy        986 ArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeu 1005
              |||||||||      |||   ||||||   ||||||   |||||||||   |||:::|||
Db       3109 CGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTA 3168

Qy       1006 GlnProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSer 1025
              |||||||||   |||      ||||||   |||   ||||||||||||||||||||||||
Db       3169 CAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGC 3228

Qy       1026 ProAlaProProAlaAspLysGlu------------------------AlaPheAlaAla 1037
              |||   |||||||||:::||||||                        ||||||
Db       3229 CCAGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACT 3288

Qy       1038 GluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProVal 1057
              |||      |||||||||   :::||||||   |||:::|||         ||||||
Db       3289 GAGGGCCA-AAGCTACCGACTGAGCCCCCACGCTGGTCATCGGCT---GCCTTCCCAT-- 3342

Qy       1058 ProProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyr 1077
              |||||||||||||||||||||   |||      |||   ||||||      |||||||||
Db       3343 CCTCCACGGGAGGTGATCAAGACTTCCACACGCGCT---GACCCT---CTCTTCTCCTAC 3396

Qy       1078 AlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuPro 1097
                 ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       3397 ACACCCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCA 3456

Qy       1098 ArgProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerVal 1117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||   |||
Db       3457 CGTCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTA 3513

Qy       1118 LeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyr 1137
              ||||||||||||:::|||||||||||||||||||||||||||||||||   ||||||:::
Db       3514 CTTGAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCAC 3573

Qy       1138 SerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAsp 1157
              |||||||||||||||   |||:::||||||:::||||||   |||||||||   :::|||
Db       3574 TCAGAGCATGCCAAG---CCCATGGGCCCTCTCACCATGGAGCTGCCCCTTGCCGTGGAC 3630

Qy       1158 ProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGln 1177
              ||||||||||||
Db       3631 CCTAAGAAGCTG------------------------------------------------ 3642

Qy       1178 AlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGly 1197
                                                                       |||
Db       3643 ---------------------------------------------------------GGG 3645

Qy       1198 ThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVal 1217
              ||||||   ||||||         ||||||||||||||||||   ||||||||||||
Db       3646 ACAGCACT-GGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCT-CCCAGTACCCGGGCT 3703

Qy       1218 ProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVal 1237
                                ::||||||||||||||||||||||   |||||||||||||||
Db       3704 GCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGC-ACGCCCGCAGACGTC 3756

Qy       1238 LeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArg 1257
              ||||||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db       3757 CTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGG 3816

Qy       1258 GlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisVal 1277
                 |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       3817 GCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTC 3876

Qy       1278 LeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1297
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       3877 CTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAGC 3936

Qy       1298 SerGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArg 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3937 TCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCGT 3996

Qy       1318 ValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPro 1337
              |||||||||   ::::::||||||||||||||||||||||||||||||||||||   |||
Db       3997 GTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CCT 4053

Qy       1338 GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGln 1357
              |||:::|||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       4054 GAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAA 4110
```

```
Qy        1358 GlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1377
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4111 GGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGGAGGCCAAGCTC 4170

Qy        1378 LeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyr 1397
               ||||||||||||||||||||||||||||||||||||   |||||||||||||||   |||
Db        4171 TTGAAGCGAGAAGGACACCACCTCCCCCACCACCACCTCGGGACCTGACTGAGACCTAC 4230

Qy        1398 LysThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeu 1414
               |||   :::            ||||||||||||||||||||||||   |||||||||:::
Db        4231 AAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGTG 4290

Qy        1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4291 GTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCGC 4350

Qy        1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4351 CGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGGC 4410

Qy        1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
               ||||||||||||||||||:::||||||   :::||||||:::||||||||||||||||||
Db        4411 ACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCGC 4470

Qy        1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
               |||:::||||||||||||||||||   ||||||   :::||||||||||||:::||||||
Db        4471 TCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCT 4530

Qy        1495 AspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThr 1514
               |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        4531 GACGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACC 4590

Qy        1515 AlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGly 1534
               :::|||   :::||||||||||||   |||||||||||||||:::|||||||||||||||
Db        4591 AGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGC 4650

Qy        1535 LysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeu 1554
               ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db        4651 AAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTG 4710

Qy        1555 ProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeu 1574
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        4711 CCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCTC 4770

Qy        1575 SerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLys 1594
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4771 CTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAAG 4830

Qy        1595 SerProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeu 1614
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4831 TCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTTG 4890

Qy        1615 LeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPro 1634
               ||||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||
Db        4891 CTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCCC 4950

Qy        1635 ThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArg 1653
               |||||||||||||||||||||||||||:::   |||||||||||||||||||||||||||
Db        4951 ACCTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCGG 5010

Qy        1654 HisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyr 1673
               ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db        5011 CACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTAC 5070

Qy        1674 ProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGln 1693
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5071 CCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAG 5130

Qy        1694 GlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeu 1713
               |||||||||||||||   |||:::||||||||||||||||||||||||||||||||||||
Db        5131 CAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTG 5190

Qy        1714 SerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAsp 1733
               |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        5191 TCACCGCGAGAGTCCTCGCTGGCCCTCAATTATTCCGCTGGCCCAAGAGGCATTATCGAC 5250

Qy        1734 LeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThr 1753
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5251 CTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACC 5310

Qy        1754 AlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSer 1773
               |||:::||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        5311 GCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGT 5370

Qy        1774 SerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSer 1793
               ||||||||||||||||||||||||||||||||||||   |||||||||   |||||||||
Db        5371 AGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCA 5430

Qy        1794 SerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
               |||||||||||||||:::|||:::||||||||||||                  ||||||
Db        5431 TCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTCC 5472

Qy        1814 IleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGln 1833
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5473 ATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAG 5532

Qy        1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSer 1853
               |||||||||:::                        |||   ||||||||||||||||||
Db        5533 AGCAGCGGGGCT------------------------GGGGGCAGCAGCCGCCCCGCCTCC 5568

Qy        1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGln 1873
               |||:::      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        5569 CACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCAG 5622

Qy        1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
               ||||||||||||||||||||||||   |||||||||::::::|||:::|||||||||
Db        5623 AGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGC 5682

Qy        1894 LysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPhe 1913
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5683 ACGCCCACGGTCCTGAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGCCACATTC 5742

Qy        1914 ProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMet 1933
               ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db        5743 CCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTACCCTCATG 5802

Qy        1934 GluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAla 1953
               ||||||||||||||||||||||||      |||||||||||||||||||||||||||
Db        5803 GAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGCCCCGTGTG 5862

Qy        1954 AspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSer 1973
               |||   |||||||||||||||   |||||||||||||||         ||||||||||||
Db        5863 GACGGTGGCCATGCCTTCCTCACCAAACCCCCGGCCCGG---------GAGCCCGCCTCC 5913

Qy        1974 SerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThr 1993
               ||||||||||||   ||||||||||||   |||   ||||||   |||   |||
Db        5914 TCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCACACAGCC 5973

Qy        1994 IleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAla 2013
               |||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        5974 ATCGCCCGCACCCCAGCAAAGAGCCTTGCACCCCACCATGCCAGTCCGGACCCGCCGGGG 6033

Qy        2014 ProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIle 2033
                  |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        6034 ---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTTTTCCATC 6090

Qy        2034 GlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSerProGluGly 2052
               |||||||||||||||||||||||||||||||||   |||:::   |||||||||:::|||
Db        6091 CAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAGCCCCGATGGG 6150

Qy        2053 ValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLys 2072
               |||||||||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db        6151 GTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGGGCTCTCCAAA 6210

Qy        2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
```

```
                  ||||||||||||:::|||||||||||||||||||||||||||   ||||||||||||
Db        6211 CCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAAGCAGCCAGGC 6270

Qy        2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSer 2112
               |||:::||||||      ||||||||||||||||||||||||||||||||||||||||||
Db        6271 CCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACTGCCCGAGAGC 6330

Qy        2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVal 2132
               |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        6331 CAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCAGAGGGTG 6390

Qy        2133 ValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPro 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6391 GTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACCCGGCACCACCCG 6450

Qy        2153 GlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPro 2172
               ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db        6451 CAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGCCAGCTGCCCT 6510

Qy        2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAla 2192
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6511 GTGCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGACCATGGCACC 6570

Qy        2193 ProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThr 2212
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        6571 CCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACCCAGCAAAACA 6630

Qy        2213 SerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThr 2232
               ||||||||||||      ||||||   |||||||||||||||||||||||||||||||||
Db        6631 TCGGTCCTGGGCAGCAGTGAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGGCATGACT 6690

Qy        2233 GluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThr 2252
               ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        6691 GAGCCAGGACATGCTCGGAGCGCTGTGTACCCACTGCTGTATCGAGACGGGGAACAGGGC 6750

Qy        2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePhe 2272
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6751 GAGCCC---AGGATGGGCTCTAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTC 6807

Qy        2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLys 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6808 AGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGATCAACAAG 6867

Qy        2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGlu 2312
               ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        6868 AAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCCTGGGACGGAA 6927

Qy        2313 IlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
               |||||||||||||||||||||||||||   ||||||||||||   |||||||||||||||
Db        6928 ATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAGCCAGGCGGTG 6987

Qy        2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6988 CAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGCACTCATGGGT 7047

Qy        2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
               |||||||||||||||||||||   |||||||||   ||||||||||||||||||||||||
Db        7048 AAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAACCCTCTGAAT 7107

Qy        2373 AlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
               ||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||||
Db        7108 GCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGGACGGAGTGAC 7167

Qy        2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
               |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7168 CACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTCTGGCAGACCTAGCAGC 7227

Qy        2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7228 CGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCCTTCTGTCTCC 7287

Qy        2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7288 TCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCGTGTGTGGGAG 7347

Qy        2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7348 GACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGATTATGAGGCTA 7407

Qy        2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
               |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        7408 CAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAGCGGGCCCCTA 7467

Qy        2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7468 GCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTCACAGTATGAG 7527

Qy        2512 ThrLeuSerAspSerGlu 2517
               ||||||||||||||||||
Db        7528 ACACTCTCGGACAGCGAG 7545

RESULT 12
US-10-764-420-1526
; Sequence 1526, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
;  APPLICANT: Lum, Pek Yee
;  APPLICANT:  Tan, Yejun
;  APPLICANT:  Dai, Hongyue
;  TITLE OF INVENTION: Methods For Determining Whether An Agent
;  TITLE OF INVENTION:  Possesses A Defined Biological Activity
;  FILE REFERENCE: ROSA122057
;  CURRENT APPLICATION NUMBER: US/10/764,420
;  CURRENT FILING DATE:  2004-01-23
;  PRIOR APPLICATION NUMBER: US 60/442,797
;  PRIOR FILING DATE: 2003-01-24
;  PRIOR APPLICATION NUMBER: US 60/474,413
;  PRIOR FILING DATE: 2003-05-30
;  NUMBER OF SEQ ID NOS: 3683
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1526
;   LENGTH: 8544
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-764-420-1526

Alignment Scores:
Pred. No.:              0               Length:          8544
Score:                  10832.50         Matches:         2149
Percent Similarity:     87.94%          Conservative:    90
Best Local Similarity:  84.41%          Mismatches:      197
Query Match:            81.97%          Indels:          117
DB:                     22              Gaps:            33

US-09-522-753-5 (1-2517) x US-10-764-420-1526 (1-8544)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db         160 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 219

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db         220 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 279

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db         280 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 339

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db         340 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGGTCTCAG 399

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||   ::::::||||||||||||||||||   ::::::
Db         400 GAGCTCCACCTGCGCCCTGAGTCCCGCACGTTCCTGCCTGAGCTGGGCAAGCCCGACATA 459
```

```
Qy      101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
            ||||||      |||||||||||||||||||||||||||||||||   ||||||||||||
Db      460 GAATTCACCGAGAGCAAGCGCCCCCGCCTGGAGCTACTACCCGATACCCTGCTGCGCCCA 519

Qy      121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
            |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db      520 TCACCCCTGCTGGCCACTGGGCAGCCGAGTGGGTCTGAAGACCTTACCAAGGACCGTAGC 579

Qy      141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
            |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db      580 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 639

Qy      161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsn---MetAspArgVal 179
            ||||||   |||   ||||||||||||||||||||||||||||||   :::|||||||||
Db      640 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACAGATTGGACCGCGTG 699

Qy      180 AspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGln 199
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      700 GACCGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAG 759

Qy      200 LeuGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIle 219
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      760 TTGGAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATA 819

Qy      220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGlu 239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      820 GAATCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGAAGAAAGCCGAA 879

Qy      240 AlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGln 259
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      880 GCCGCACACCGGATCCTAGAAGGCCTGGGGCCCCAGGTGGAGCTGCCTCTGTACAACCAG 939

Qy      260 ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys 279
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      940 CCGTCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAG 999

Qy      280 LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys 299
            |||||||||||||||||||||||||||||||||||||||||||||:::|||:::||||||
Db     1000 CTGATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGGAACAGCGCTTCTGC 1059

Qy      300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsn 319
            |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db     1060 CAGCGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAAT 1119

Qy      320 ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1120 CCGCGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAG 1179

Qy      340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer 359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1180 ATCCGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGT 1239

Qy      360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1240 GGGCTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTG 1299

Qy      380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
            ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db     1300 TCTGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCC--GCCATGTTG 1357

Qy      400 TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro 419
            ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db     1358 TACGACGC-GACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCC 1416

Qy      420 MetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThr 439
            |||||||||||||||||||||||||||   |||||||||||||||||||||::::::|||
Db     1417 ATGAAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACC 1476

Qy      440 PheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu 459
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1477 TTCCGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAG 1536

Qy      460 ArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyr 479
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1537 AGAAAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTAC 1596

Qy      480 LysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGln 499
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1597 AAGAGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAG 1656

Qy      500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
            ||||||||||||||||||||||||            |||   ||||||||||||||||||
Db     1657 CAACAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAG 1704

Qy      520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
            |||:::||||||||||||||||||||||||:::|||||||||||||||   :::   |||
Db     1705 AAGGAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAG 1764

Qy      540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
            |||:::||||||:::|||   |||||||||||||||||||||||||||||||||||||||
Db     1765 AACGAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACGAT 1824

Qy      560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1825 GAGAAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAA 1884

Qy      580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
            ||||||||||||||||||||||||||||||||||||   ||||||      |||||||||
Db     1885 GGCCGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAG 1944

Qy      600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
            ||||||:::||||||||||||||||||:::||||||||||||||||||||||||||||||
Db     1945 CAAAGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAG 2004

Qy      620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2005 ATGGAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGC 2064

Qy      640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2065 ATGGTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2124

Qy      660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2125 CAGAACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCT 2184

Qy      680 ArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProVal 699
            ||||||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db     2185 CGGAGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCC 2244

Qy      700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
               |||||||||||||||||||||||||||   |||   ||||||||||||:::   |||
Db     2245 GCTGAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAG 2304

Qy      720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGly 738
            ||||||||||||      ||||||||||||||||||||||||   |||||||||||||||
Db     2305 GAGGCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGC 2364

Qy      739 ProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAla 758
            ||||||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db     2365 CCAGCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCC 2424

Qy      759 AlaLysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAsp 776
               |||||||||         ||||||||||||      |||      |||   |||
Db     2425 ACGAAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCC 2475

Qy      777 GlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluPro 796
                  ||||||   |||      ||||||         :::   ||||||   ||||||
Db     2476 ACCCAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGCAGCCCCTGCTGAGCCC 2529

Qy      797 ThrProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSer 816
            :::|||      :::|||:::|||   |||:::|||   |||:::   |||||||||:::
Db     2530 TCCCCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTCC---CCATCACCTGCC 2586

Qy      817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
            |||||||||      |||   |||:::|||:::|||         |||   ||||||
Db       2587 GCACCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAG 2646

Qy        837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
                 |||:::   :::|||||||||      |||   |||         :::|||   |||
Db       2647 ACAGAAGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGGA 2697

Qy        857 LysAlaGluGluPro---------------------------ValLysSerGluCysThr 867
              |||   |||||||||                           |||||||||:::
Db       2698 AAGCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTAAAGAGTGACCACAAG 2757

Qy        868 GluGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThr 886
              ||||||   ||||||   |||      |||   |||      ||||||   |||
Db       2758 GAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTG 2817

Qy        887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThr--- 905
              :::|||      ||||||   |||      |||   ||||||      |||   |||
Db       2818 TCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAGGCAGCTGTGACCAAG 2871

Qy        906 AlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGlu 925
                    ||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       2872 GGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACCTGCAGTGCCGATGAG 2931

Qy        926 ValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeu 945
              |||||||||   |||||||||||||||   ||||||||||||||||||||||||||||||
Db       2932 GTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTC 2991

Qy        946 ThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGln 965
              ||||||   |||||||||||||||:::   ||||||||||||||||||||||||||||||
Db       2992 ACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAG 3051

Qy        966 LeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProPro 985
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       3052 CTGAAGCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAGGTCCATGAGCCCCCC 3108

Qy        986 ArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeu 1005
              |||||||||      |||   ||||||   ||||||   |||||||||   |||:::|||
Db       3109 CGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTA 3168

Qy       1006 GlnProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSer 1025
              |||||||||   |||      ||||||   |||   ||||||||||||||||||||||||
Db       3169 CAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGC 3228

Qy       1026 ProAlaProProAlaAspLysGlu------------------------AlaPheAlaAla 1037
              |||   |||||||||:::||||||                        ||||||
Db       3229 CCAGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACT 3288

Qy       1038 GluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProVal 1057
              |||      |||||||||   :::||||||   |||:::|||         ||||||
Db       3289 GAGGGCCA-AAGCTACCGACTGAGCCCCCACGCTGGTCATCGGCT---GCCTTCCCAT-- 3342

Qy       1058 ProProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyr 1077
              |||||||||||||||||||||   |||      |||   ||||||      |||||||||
Db       3343 CCTCCACGGGAGGTGATCAAGACTTCCACACGCGCT---GACCCT---CTCTTCTCCTAC 3396

Qy       1078 AlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuPro 1097
                 ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       3397 ACACCCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCA 3456

Qy       1098 ArgProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerVal 1117
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||   |||
Db       3457 CGTCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTA 3513

Qy       1118 LeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyr 1137
              ||||||||||||:::|||||||||||||||||||||||||||||||||   ||||||:::
Db       3514 CTTGAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCAC 3573

Qy       1138 SerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAsp 1157
              |||||||||||||||   |||:::||||||:::||||||   |||||||||   :::|||
Db       3574 TCAGAGCATGCCAAG---CCCATGGGCCCTCTCACCATGGAGCTGCCCCTTGCCGTGGAC 3630

Qy       1158 ProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGln 1177
              ||||||||||||
Db       3631 CCTAAGAAGCTG------------------------------------------------ 3642

Qy       1178 AlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGly 1197
                                                                       |||
Db       3643 ---------------------------------------------------------GGG 3645

Qy       1198 ThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVal 1217
              ||||||   ||||||         ||||||||||||||||||   ||||||||||||
Db       3646 ACAGCACT-GGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCT-CCCAGTACCCGGGCT 3703

Qy       1218 ProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVal 1237
                                :::|||||||||||||||||||||   |||||||||||||||
Db       3704 GCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGC-ACGCCCGCAGACGTC 3756

Qy       1238 LeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArg 1257
              ||||||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db       3757 CTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGG 3816

Qy       1258 GlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisVal 1277
                 |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       3817 GCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTC 3876

Qy       1278 LeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1297
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       3877 CTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAGC 3936

Qy       1298 SerGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArg 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3937 TCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCGT 3996

Qy       1318 ValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPro 1337
              |||||||||   ::::::||||||||||||||||||||||||||||||||||||   |||
Db       3997 GTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CCT 4053

Qy       1338 GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGln 1357
              |||:::|||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       4054 GAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAA 4110

Qy       1358 GlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4111 GGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGGAGGCCAAGCTC 4170

Qy       1378 LeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyr 1397
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||   |||
Db       4171 TTGAAGCGAGAAGGGACACCACCTCCCCCACCACCACCTCGGGACCTGACTGAGACCTAC 4230

Qy       1398 LysThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeu 1414
              |||   :::               ||||||||||||||||||||||||   |||||||||:::
Db       4231 AAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGTG 4290

Qy       1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4291 GTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCGC 4350

Qy       1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4351 CGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGGC 4410

Qy       1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
              ||||||||||||||||||:::||||||   :::||||||:::||||||||||||||||||
Db       4411 ACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCGC 4470

Qy       1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
              |||:::||||||||||||||||||   ||||||   :::||||||||||||:::||||||
Db       4471 TCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCT 4530

Qy       1495 AspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThr 1514
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       4531 GACGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACC 4590

Qy       1515 AlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGly 1534
              :::|||   :::||||||||||||   |||||||||||||||:::|||||||||||||||
Db       4591 AGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGC 4650
```

```
Qy       1535 LysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeu 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db       4651 AAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTG 4710

Qy       1555 ProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeu 1574
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       4711 CCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCTC 4770

Qy       1575 SerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLys 1594
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4771 CTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAAG 4830

Qy       1595 SerProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeu 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4831 TCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTTG 4890

Qy       1615 LeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPro 1634
              ||||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||
Db       4891 CTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCCC 4950

Qy       1635 ThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArg 1653
              |||||||||||||||||||||||||||:::   |||||||||||||||||||||||||||
Db       4951 ACCTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCGG 5010

Qy       1654 HisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyr 1673
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       5011 CACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTAC 5070

Qy       1674 ProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGln 1693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5071 CCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAG 5130

Qy       1694 GlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeu 1713
              |||||||||||||||   |||:::||||||||||||||||||||||||||||||||||||
Db       5131 CAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTG 5190

Qy       1714 SerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAsp 1733
              |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db       5191 TCACCGCGAGAGTCCTCGCTGGCCCTCAATTATTCCGCTGGCCCAAGAGGCATTATCGAC 5250

Qy       1734 LeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThr 1753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5251 CTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACC 5310

Qy       1754 AlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSer 1773
              |||:::||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db       5311 GCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGT 5370

Qy       1774 SerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSer 1793
              ||||||||||||||||||||||||||||||||||||   |||||||||   |||||||||
Db       5371 AGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCA 5430

Qy       1794 SerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
              |||||||||||||||:::|||:::||||||||||||                  ||||||
Db       5431 TCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTCC 5472

Qy       1814 IleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGln 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5473 ATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAG 5532

Qy       1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSer 1853
              |||||||||:::                        |||   ||||||||||||||||||
Db       5533 AGCAGCGGGGCT------------------------GGGGGCAGCAGCCGCCCCGCCTCC 5568

Qy       1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGln 1873
              |||:::      ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5569 CACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCAG 5622

Qy       1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
              ||||||||||||||||||||||||   |||||||||::::::|||:::|||||||||
Db       5623 AGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGC 5682

Qy       1894 LysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPhe 1913
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5683 ACGCCCACGGTCCTGAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGCCACATTC 5742

Qy       1914 ProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMet 1933
              ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db       5743 CCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTACCCTCATG 5802

Qy       1934 GluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAla 1953
              ||||||||||||||||||||||||      |||||||||||||||||||||||||||
Db       5803 GAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGCCCCGTGTG 5862

Qy       1954 AspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSer 1973
              |||   |||||||||||||||   |||||||||||||||         ||||||||||||
Db       5863 GACGGTGGCCATGCCTTCCTCACCAAACCCCCGGCCCGG---------GAGCCCGCCTCC 5913

Qy       1974 SerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThr 1993
              ||||||||||||   ||||||||||||   |||   ||||||   |||   |||
Db       5914 TCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCACACAGCC 5973

Qy       1994 IleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAla 2013
              |||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db       5974 ATCGCCCGCACCCCAGCAAAGAGCCTTGCACCCCACCATGCCAGTCCGGACCCGCCGGGG 6033

Qy       2014 ProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIle 2033
                 |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       6034 ---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTTTTCCATC 6090

Qy       2034 GlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSerProGluGly 2052
              |||||||||||||||||||||||||||||||||   |||:::   |||||||||:::|||
Db       6091 CAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAGCCCCGATGGG 6150

Qy       2053 ValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLys 2072
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db       6151 GTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGGGCTCTCCAAA 6210

Qy       2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
                 ||||||||||||:::|||||||||||||||||||||||||||   ||||||||||||
Db       6211 CCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAAGCAGCCAGGC 6270

Qy       2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSer 2112
              |||:::||||||      ||||||||||||||||||||||||||||||||||||||||||
Db       6271 CCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACTGCCCGAGAGC 6330

Qy       2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVal 2132
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db       6331 CAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCAGAGGGTG 6390

Qy       2133 ValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPro 2152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6391 GTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACCCGGCACCACCCG 6450

Qy       2153 GlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPro 2172
              ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db       6451 CAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGCCAGCTGCCCT 6510

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAla 2192
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6511 GTGCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGACCATGGCACC 6570

Qy       2193 ProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThr 2212
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       6571 CCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACCCAGCAAAACA 6630

Qy       2213 SerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThr 2232
              ||||||||||||      ||||||   |||||||||||||||||||||||||||||||||
Db       6631 TCGGTCCTGGGCAGCAGTGAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGGCATGACT 6690

Qy       2233 GluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThr 2252
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       6691 GAGCCAGGACATGCTCGGAGCGCTGTGTACCCACTGCTGTATCGAGACGGGGAACAGGGC 6750

Qy       2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePhe 2272
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6751 GAGCCC---AGGATGGGCTCTAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTC 6807

Qy        2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLys 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6808 AGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGATCAACAAG 6867

Qy        2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGlu 2312
               ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        6868 AAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCCTGGGACGGAA 6927

Qy        2313 IlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
               |||||||||||||||||||||||||||   ||||||||||||   |||||||||||||||
Db        6928 ATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAGCCAGGCGGTG 6987

Qy        2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6988 CAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGCACTCATGGGT 7047

Qy        2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
               |||||||||||||||||||||   |||||||||   ||||||||||||||||||||||||
Db        7048 AAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAACCCTCTGAAT 7107

Qy        2373 AlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
               ||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||||
Db        7108 GCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGGACGGAGTGAC 7167

Qy        2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
               |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7168 CACGCACTCACCTCGCCAGGTGGAGGTGGAAAGCCAAGGTCTCTGGCAGACCTAGCAGC 7227

Qy        2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7228 CGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCCTTCTGTCTCC 7287

Qy        2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7288 TCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCGTGTGTGGGAG 7347

Qy        2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7348 GACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGATTATGAGGCTA 7407

Qy        2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
               |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        7408 CAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAGCGGGCCCCTA 7467

Qy        2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7468 GCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTCACAGTATGAG 7527

Qy        2512 ThrLeuSerAspSerGlu 2517
               ||||||||||||||||||
Db        7528 ACACTCTCGGACAGCGAG 7545

RESULT 13
US-10-087-192-650
; Sequence 650, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 650
```

```
;   LENGTH: 7534
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-087-192-650

Alignment Scores:
Pred. No.:                 0               Length:           7534
Score:                     9691.50         Matches:          1893
Percent Similarity:        87.70%          Conservative:     82
Best Local Similarity:     84.06%          Mismatches:       171
Query Match:               73.34%          Indels:           106
DB:                        14              Gaps:             28

US-09-522-753-5 (1-2517) x US-10-087-192-650 (1-7534)

Qy         295 LysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGlu 314
               :::|||:::|||||||||||||||||||||||||||||||||   |||||||||||||||
Db           2 GAACAGCGCTTCTGCCAGCGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAG 61

Qy         315 ArgIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGlu 334
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 CGCATAGAGAACAATCCGCGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAG 121

Qy         335 LysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgVal 354
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 AAACAGTTCCCGGAGATCCGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTG 181

Qy         355 GlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGlu 374
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 GGCCAGCGTGGCAGTGGGCTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAG 241

Qy         375 IleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaVal 394
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 ATCATTGATGGCTTGTCTGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTG 301

Qy         395 IleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGly 414
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         302 ATCCCGCCCATGTTGTACGACGCGGACCAGCAGAGGATCAAGTTCATCAACATGAATGGA 361

Qy         415 LeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGlu 434
               ||||||   |||||||||||||||||||||||||||||||||   |||||||||||||||
Db         362 CTCATGGATGACCCCATGAAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAG 421

Qy         435 GlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIle 454
               ||||||::::::||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CAGGAGAGGGACACCTTCCGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATT 481

Qy         455 AlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLys 474
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 GCCTCATTCCTGGAGAGAAAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAG 541

Qy         475 LysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGln 494
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 AAGAATGAAAATTACAAGAGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAG 601

Qy         495 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArg 514
               |||||||||||||||||||||||||||||||||||||||            |||   |||
Db         602 CAGCAGCAGCAGCAGCAACAACAGCAGCAGCAGCAGCAG------------ATGGCACGG 649

Qy         515 SerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGlu 534
               ||||||||||||||||||:::||||||||||||||||||||||||:::||||||||||||
Db         650 AGCAGCCAGGAGGAGAAGGAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAG 709

Qy         535 LysProGluValGluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSer 554
               |||   :::   ||||||:::||||||:::|||   ||||||||||||||||||||||||
Db         710 AAGCAGGATGCGGAGAACGAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCT 769

Qy         555 GlyGluAspAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSer 574
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         770 GGCGAGGACAACGATGAGAAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGC 829

Qy         575 GlnGlyArgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGlu 594
               |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        830 CAAGGCCGCCGCAAAGGCCGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAG 889

Qy        595 AlaIleThrProGlnGlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArg 614
                    |||||||||||||||:::||||||||||||||||||:::|||||||||||||||
Db        890 ACAGCCACCCCACAGCAAAGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGC 949

Qy        615 TrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrp 634
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        950 TGGACTGAGGAAGAGATGGAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGG 1009

Qy        635 SerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPhe 654
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1010 TCAGCCATTGCCCGCATGGTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTC 1069

Qy        655 AsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGlu 674
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1070 AACTACAAGAAGAGGCAGAACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAG 1129

Qy        675 LysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAla 694
              |||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db       1130 AAGGAGAGGAACGCTCGGAGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACA 1189

Qy        695 AlaPheProProValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGlu 714
              ||||||||||||      |||||||||||||||||||||||||||   |||   ||||||
Db       1190 GCCTTCCCACCTGCCGCTGAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAG 1249

Qy        715 GluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArg--- 733
              ||||||:::   |||||||||||||||      ||||||||||||||||||||||||
Db       1250 GAAGAGCTGGCGGAGGAGGCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTT 1309

Qy        734 GlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSer 753
              |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
Db       1310 GGGGAGTGCAGTGGCCCAGCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCC 1369

Qy        754 ProHisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProProAlaThr--- 772
              |||   :::||||||   |||||||||         ||||||||||||      |||
Db       1370 CCGCGTTCAGAAGCCACGAAGGACACT---------GGGCCTAAACCCACTGGCACTGAA 1420

Qy        773 ---LeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThrSerArg 791
                 |||   |||         ||||||   |||      ||||||         :::
Db       1421 GCATTGCCCGCTGCCACCCAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGCA 1474

Qy        792 AlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProProAla 811
              ||||||   ||||||:::|||      :::|||:::|||   |||:::|||   |||:::
Db       1475 GCCCCTGCTGAGCCCTCCCCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTCC 1534

Qy        812 ProProSerProSerAlaProProProValValProLysGluGluLysGluGluGluThr 831
                 |||||||||:::|||||||||      |||   |||:::|||:::|||
Db       1535 ---CCATCACCTGCCGCACCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCT 1591

Qy        832 AlaAlaAlaProProValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeu 851
              |||   ||||||      |||:::   :::|||||||||      |||   |||
Db       1592 GCTCCAGCTCCCCAGACAGAAGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG------ 1645

Qy        852 AlaValAspThrGlyLysAlaGluGluPro---------------------------Val 862
                 :::|||   ||||||   |||||||||                           |||
Db       1646 ---ATCGATGTGGGAAAGCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTA 1702

Qy        863 LysSerGluCysThrGluGluAlaGluGluGlyProAla---LysGlyLysAspAlaGlu 881
              ||||||:::      ||||||   ||||||   |||      |||   |||      |||
Db       1703 AAGAGTGACCACAAGGAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAG 1762

Qy        882 AlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGly 901
              |||   |||      :::|||      ||||||   |||      |||   ||||||
Db       1763 GCCATTGAAACTGTGTCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAG 1816

Qy        902 ArgAlaThrThr---AlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
                 |||   |||         ||||||||||||   ||||||||||||||||||||||||
Db       1817 GCAGCTGTGACCAAGGGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACC 1876

Qy        921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
              ||||||||||||||||||||||||   |||||||||||||||   |||||||||||||||
Db       1877 TGCAGTGCCGATGAGGTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCA 1936
```

```
Qy        941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
              |||||||||||||||||||||   |||||||||||||||:::   |||||||||||||||
Db       1937 AGGCCCAGCCTCCTCACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCG 1996

Qy        961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db       1997 CTGGACCTGAAGCAGCTGAAGCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAG 2053

Qy        981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
              ||||||||||||||||||||||||      |||   ||||||   ||||||   ||||||
Db       2054 GTCCATGAGCCCCCCCGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCA 2113

Qy       1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
              |||   |||:::||||||||||||   |||      ||||||   |||   |||||||||
Db       2114 CCCACGCAGCACCTACAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGT 2173

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGlu--------------------- 1033
              ||||||||||||||||||   |||||||||:::||||||
Db       2174 GGCAAGTCCCGCAGCCCAGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTT 2233

Qy       1034 ---AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGly 1052
                 ||||||      |||   |||||||||   :::||||||   |||:::||||||
Db       2234 CCGGCTTTCCCAACTGAGGGCCCAAAGCTACCGACTGAGCCCCCACGCTGGTCATCGGGC 2293

Qy       1053 LeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPro 1072
              ||||||||||||:::|||||||||||||||||||||   ||||||||||||   ||||||
Db       2294 CTGCCCTTCCCCATCCCTCCACGGGAGGTGATCAAGACTTCCCCACACGCCGCTGACCCC 2353

Qy       1073 SerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAla 1092
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||:::|||
Db       2354 TCTGCCTTCTCCTACACACCCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCC 2413

Qy       1093 ArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAla 1112
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       2414 CGGCCCGTCCTGCCACGTCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCC 2470

Qy       1113 LysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGln 1132
              |||||||||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db       2471 AAGCATCCCGGCGTACTTGAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAG 2530

Qy       1133 LeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLeu 1152
              |||   ||||||:::|||||||||||||||||||||:::||||||:::||||||||||||
Db       2531 CTTCGTGTGCCTCACTCAGAGCATGCCAAGGCCCCCATGGGCCCTCTCACCATGGGGCTG 2590

Qy       1153 ProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeu 1172
              ||||||   :::|||||||||||||||
Db       2591 CCCCTTGCCGTGGACCCTAAGAAGCTG--------------------------------- 2617

Qy       1173 SerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAla 1192

Db       2617 ------------------------------------------------------------ 2617

Qy       1193 SerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIle 1212
                          ||||||||||||||||||         ||||||||||||||||||:::
Db       2618 ------------GGGACAGCACTGGGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCTC 2665

Qy       1213 ProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGly 1232
              ||||||||||||                     :::||||||||||||||||||||||||
Db       2666 CCCAGTACCCGGGCTGCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGGC 2719

Qy       1233 ThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPro 1252
              |||||||||||||||||||||||||||||||||:::||||||:::|||||||||||||||
Db       2720 ACGCCCGCAGACGTCCTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCA 2779

Qy       1253 SerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGly 1272
              |||||||||||||||   |||||||||:::||||||||||||||||||||||||||||||
Db       2780 AGTCGCCTTGACCGGGCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGC 2839

Qy       1273 LysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGlu 1292
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       2840 AAGAAAGGCCACGTCCTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAG 2899
```

```
Qy          1293 AspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAsp 1312
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2900 GATGGAAGGAGCAGCTCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGAC 2959

Qy          1313 MetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGly 1332
                 ||||||||||||||||||||||||   ::::::|||||||||||||||||||||||||||
Db          2960 ATGATGGAGGGCCGTGTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGC 3019

Qy          1333 ArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleArg 1352
                 |||||||||   ||||||:::|||||||||   |||||||||||||||||||||||||||
Db          3020 CGCGCCATC---CCTGAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGA 3073

Qy          1353 GlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArg 1372
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3074 GGCTCCATCACGCAAGGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGG 3133

Qy          1373 ArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAsp 1392
                 |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db          3134 CGGGAGGCCAAGCTCTTGAAGCGAGAAGGACACCACCACCCCCACCACCACCTCGGGAC 3193

Qy          1393 LeuThrGluAlaTyrLysThrGln---------AlaLeuGlyProLeuLysLeuLysPro 1409
                 |||||||||   ||||||   :::            ||||||||||||||||||||||||
Db          3194 CTGACTGAGACCTACAAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCG 3253

Qy          1410 AlaHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIlePro 1429
                    |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db          3254 ACTCACGAGGGTGTGGTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCG 3313

Qy          1430 ArgGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGly 1449
                 |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db          3314 AGAGAGGAGCTGCGCCGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGT 3373

Qy          1450 SerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLys 1469
                 |||||||||||||||||||||||||||||||||:::||||||   :::||||||:::|||
Db          3374 TCCATCACCCAGGGCACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAG 3433

Qy          1470 LysHisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisPro 1489
                 ||||||||||||||||||:::||||||||||||||||||   ||||||   :::||||||
Db          3434 AAACACGACGTGCGCTCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCG 3493

Qy          1490 LeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLys 1509
                 ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3494 CTGGACATAATGGCTGATGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAG 3553

Qy          1510 SerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIle 1529
                 ||||||   ||||||:::|||   :::||||||||||||   |||||||||||||||:::
Db          3554 AGCCGGTCAGGGACCAGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTC 3613

Qy          1530 ValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaPro 1549
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3614 GTGCCTGAACTGGGCAAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCC 3673

Qy          1550 PheAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeu 1569
                 |||      |||||||||||||||||||||||||||   |||||||||||||||||||||
Db          3674 TTCACCAGTCACCTGCCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTT 3733

Qy          1570 GlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrPro 1589
                 |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db          3734 CAGGAAGGCAGCCTCCTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCC 3793

Qy          1590 ArgGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSer 1609
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3794 CGGGAGATCGCCAAGTCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCC 3853

Qy          1610 ProTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIlePro 1629
                 |||||||||||||||||||||||||||:::||||||||||||||||||   |||||||||
Db          3854 CCCTATGAGCACTTGCTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCA 3913

Qy          1630 LeuAlaPheAspProThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAla 1648
                 ||||||||||||||||||||||||||||||||||||||||||:::   ||||||||||||
Db          3914 TTGGCCTTTGACCCCACCTCCATACCCCGAGGATCCCTCTGGAAGCAGCAGCCGCAGCC 3973

Qy          1649 TyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyr 1668
                 |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db          3974 TACTACCTGCCCCGGCACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTAC 4033

Qy          1669 LeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAsp 1688
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4034 CTCATCCGCGGCTACCCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGAC 4093

Qy          1689 TyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAsp 1708
                 ||||||||||||||||||||||||||||||   |||:::|||||||||||||||||||||
Db          4094 TACATCACCTCGCAGCAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGAC 4153

Qy          1709 MetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyPro 1728
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4154 ATGCTGAGGGGTCTGTCACCGCGAGAGTCCTCGCTGGCCCTCAATTATGCCGCTGGCCCA 4213

Qy          1729 ArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrPro 1748
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4214 AGAGGCATTATCGACCTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCA 4273

Qy          1749 GlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPhe 1768
                 ||||||||||||||||||:::||||||||||||||||||||||||||||||   ||||||
Db          4274 GGCACCCCTGCCACCGCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTC 4333

Qy          1769 SerSerArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysPro 1788
                 |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db          4334 AGCAGCCGCCACAGTAGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCA 4393

Qy          1789 ThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArg 1808
                 |||   ||||||||||||||||||||||||:::|||:::||||||||||||
Db          4394 ACTGCCACATCTTCATCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC--------- 4444

Qy          1809 GluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArg 1828
                          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4445 ---------AAGTCCATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGA 4495

Qy          1829 ProGlyThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSer 1848
                 |||||||||||||||||||||||||||:::                     |||   |||
Db          4496 CCTGGTACGGAGCAGAGCAGCGGGGCT---------------------GGGGGCAGC 4531

Qy          1849 SerArgProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGln 1868
                 ||||||||||||||||||:::      |||||||||||||||||||||||||||||||||
Db          4532 AGCCGCCCCGCCTCCCACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAG 4585

Qy          1869 AspAlaLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThr 1888
                 |||||||||||||||||||||||||||||||||||||||   |||||||||::::::|||
Db          4586 GACGCCTTGCAGCAGAGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACC 4645

Qy          1889 Ala-ValGluProSerLysProThrValLeuArgSerThrSerThrSerSerProValAr 1908
                 ::: ||||||||       |||||||||||||||||||||||||||||||||||||||||
Db          4646 TCCCGTGGAACCCGGCACGCCCACGGTCCTGAGGTCCACCTCCACCTCTTCGCCTGTCCG 4705

Qy          1908 gProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyVa 1928
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db          4706 CCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGT 4765

Qy          1928 lTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgPr 1948
                 ||||||||||||||||||||||||||||||||||||||||      ||||||||||||||
Db          4766 CTACCCTACCCTCATGGAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCC 4825

Qy          1948 oGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGl 1968
                 |||||||||||||   |||   |||||||||||||||   |||||||||||||||
Db          4826 CGAGCGGCCCCGTGTGGACGCTGGCCATGCCTTCCTCACCAAACCCCCGGCCCGG----- 4880

Qy          1968 yLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProProVa 1988
                     ||||||||||||||||||||||||   ||||||||||||   |||   ||||||
Db          4881 ----GAGCCCGCCTCCTCACCCAGCAAGAGCTCCGAGCCCCGATCCTAGCACCCCCCAG 4936

Qy          1988 lSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSe 2008
                  |||   |||      ||||||||||||||||||||:::|||||||||||||||||||||
Db          4937 CTCCAGCCACACAGCCATCGCCCGCACCCCAGCAAAGAGCCTTGCACCCCACCATGCCAG 4996

Qy          2008 rProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSe 2028
                 ||||||||||||||||    |||   |||||||||||    |||||||||||||||||||
Db       4997 TCCGGACCCGCCGGCG---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAG 5053

Qy       2028 rLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSe 2047
              |||||||||||||||||||||||||||||||||||||||||||||||||   |||:::
Db       5054 TAAACCCTTTTCCATCCAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGG 5113

Qy       2047 rTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAs 2067
               |||||||||:::||||||||||||:::||||||||||||||||||||||||||||||||
Db       5114 CTACAGCCCCGATGGGGTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGA 5173

Qy       2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuAr 2087
              ||||||||||   |||   ||||||||||||:::||||||||||||||||||||||||||
Db       5174 CAAGGGGCTCTCCAAACCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCG 5233

Qy       2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107
              |   |||||||||||||||:::||||||      ||||||||||||||||||||||||||
Db       5234 GCACAAGCAGCCAGGCCCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCG 5293

Qy       2107 gProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLy 2127
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:::||
Db       5294 GCCACTGCCCGAGAGCCAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAA 5353

Qy       2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5354 AGGTCACCAGAGGGTGGTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTA 5413

Qy       2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
              ||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db       5414 CACGCGCCACCACCCGCAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCC 5473

Qy       2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5474 CGGAGCCAGCTGCCCTGTCCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACC 5533

Qy       2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerPr 2207
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       5534 CCCCGACCATGGCACCCCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCC 5593

Qy       2207 oGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerPr 2227
              |||||||:::||||||||||||||||||      ||||||   |||||||||||||||||
Db       5594 AGAACCCAGCAAAACATCGGTCCTGGGCAGCAGTGAGGATGCCATTGAGCCTGTGTCCCC 5653

Qy       2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db       5654 ACCAGAGGGCATGACTGAGCCAGGACATGCTCGGAGCGCTGTGTACCCACTGCTGTATCG 5713

Qy       2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGl 2267
              |||||||||||||   ||||||   |||||||||||||||||||||||||||||||||||
Db       5714 AGACGGGGAACAGGGCGAGCCC---AGGATGGGCTCTAAGTCTCCAGGCAACACCAGCCA 5770

Qy       2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLy 2287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5771 GCCGCCAGCCTTCTTCAGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAA 5830

Qy       2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5831 GCAGGAGATCAACAAGAAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGG 5890

Qy       2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327
               ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db       5891 CCAGCCTGGGACGGAAATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTG 5950

Qy       2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5951 TAGAAGCCAGGCGGTGCAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAG 6010

Qy       2347 gLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAl 2367
              |||||||||||||||||||||||||||||||||||||   |||||||||   ||||||||
Db       6011 AAAGGCACTCATGGGTAAATATGATCAGTGGAAGAGCCCCCGCCGCTCGGCGCCAATGC 6070

Qy       2367 aPheAsnProLeuAsnAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAl 2386
              ||||||||||||||||||||||||||||||||||   ||||||||||||||||||   ||
Db       6071 TTTTAACCCTCTGAATGCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGC 6130
```

```
Qy       2386 aAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSe 2406
              |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       6131 TGACGGACGGAGTGACCACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTC 6190

Qy       2406 rGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspAr 2426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6191 TGGCAGACCTAGCAGCCGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCG 6250

Qy       2426 gProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuTh 2446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6251 ACCCCCTTCTGTCTCCTCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCAC 6310

Qy       2446 rAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnPr 2466
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6311 CAACCGTGTGTGGGAGGACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCC 6370

Qy       2466 oLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAl 2486
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||   ||
Db       6371 TTTGATTATGAGGCTACAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGC 6430

Qy       2486 aGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLe 2506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6431 AGGCAGCGGGCCCCTAGCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCT 6490

Qy       2506 uCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||||||||
Db       6491 GTGTTCACAGTATGAGACACTCTCGGACAGCGAG 6524

RESULT 14
US-10-146-473-14
; Sequence 14, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew
;  APPLICANT:  Gout, Ivan
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Gure, Ali
;  APPLICANT:  Chen, Yao-Tseng
;  APPLICANT:  Old, Lloyd
;  TITLE OF INVENTION: Breast Cancer Antigens
;  FILE REFERENCE: L00461/70130(JRV)
;  CURRENT APPLICATION NUMBER: US/10/146,473
;  CURRENT FILING DATE:  2002-05-15
;  PRIOR APPLICATION NUMBER: US 60/291,150
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 2930
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-146-473-14

Alignment Scores:
Pred. No.:                7.66e-264      Length:        2930
Score:                    4612.00        Matches:       873
Percent Similarity:       94.38%         Conservative:  1
Best Local Similarity:    94.28%         Mismatches:    6
Query Match:              34.90%         Indels:        46
DB:                       16             Gaps:          1

US-09-522-753-5 (1-2517) x US-10-146-473-14 (1-2930)

Qy       1592 IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyr 1611
              :::         ||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CTTGGCTCACCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTAT 182

Qy       1612 GluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCC 242

Qy       1632 PheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeu 1651
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          243 TTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTG 302

Qy         1652 ProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArg 1671
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          303 CCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGC 362

Qy         1672 GlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThr 1691
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          363 GGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACC 422

Qy         1692 SerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArg 1711
                |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db          423 TCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGG 482

Qy         1712 GlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIle 1731
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          483 GGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATC 542

Qy         1732 IleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrPro 1751
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          543 ATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCA 602

Qy         1752 AlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg 1771
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          603 GCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGC 662

Qy         1772 HisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThr 1791
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          663 CACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACG 722

Qy         1792 SerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGlu 1811
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          723 TCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAA 782

Qy         1812 LysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThr 1831
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          783 AAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACA 842

Qy         1832 GluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgPro 1851
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          843 GAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCC 902

Qy         1852 AlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeu 1871
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          903 GCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTC 962

Qy         1872 GlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGlu 1891
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          963 CAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAG 1022

Qy         1892 ProSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAla 1911
                ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1023 CCCAGCACGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCC 1082

Qy         1912 ThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThr 1931
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1083 ACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACC 1142

Qy         1932 LeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgPro 1951
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1143 CTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCC 1202

Qy         1952 ArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluPro 1971
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1203 CGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCC 1262

Qy         1972 AlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHis 1991
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1263 GCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCAC 1322

Qy         1992 AlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspPro 2011
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         1323 GCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCG 1382

Qy         2012 ProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPhe 2031
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1383 CCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTT 1442

Qy         2032 SerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGlu 2051
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1443 TCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAA 1502

Qy         2052 GlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuPro 2071
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1503 GGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCC 1562

Qy         2072 LysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnPro 2091
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1563 AAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCA 1622

Qy         2092 GlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGlu 2111
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1623 GGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAG 1682

Qy         2112 SerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArg 2131
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1683 AGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGG 1742

Qy         2132 ValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHis 2151
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1743 GTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCAC 1802

Qy         2152 ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCys 2171
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1803 CCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGC 1862

Qy         2172 ProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGly 2191
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1863 CCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGT 1922

Qy         2192 AlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLys 2211
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1923 GCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAG 1982

Qy         2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1983 ACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATG 2042

Qy         2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2043 ACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAG 2102

Qy         2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2103 ACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTC 2162

Qy         2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2163 TTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAAC 2222

Qy         2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2223 AAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACG 2282

Qy         2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2283 GAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCG 2342

Qy         2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2343 GTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATG 2402

Qy         2352 GlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeu 2371

Db         2402 ------------------------------------------------------------ 2402
```

```
Qy        2372 AsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391

Db        2402 ------------------------------------------------------------ 2402

Qy        2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
                                 ||||||||||||||||||||||||||||||||||||||||||
Db        2403 ------------------GGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGC 2444

Qy        2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2445 CGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCC 2504

Qy        2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2505 TCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAG 2564

Qy        2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2565 GACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTG 2624

Qy        2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2625 CAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTC 2684

Qy        2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2685 GCTGGCGCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAG 2744

Qy        2512 ThrLeuSerAspSerGlu 2517
               ||||||||||||||||||
Db        2745 ACACTCTCCGACAGCGAG 2762

RESULT 15
US-10-174-014-13
; Sequence 13, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
;  APPLICANT: C. Frank Bennett
;  APPLICANT:  Susan M. Freier
;  APPLICANT:  Kenneth W. Dobie
;  TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
;  FILE REFERENCE: PTS-0012
;  CURRENT APPLICATION NUMBER: US/10/174,014
;  CURRENT FILING DATE:  2002-06-17
;  NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 13
;   LENGTH: 2930
;   TYPE: DNA
;   ORGANISM: H. sapiens
;   FEATURE:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (456)...(2765)
US-10-174-014-13

Alignment Scores:
Pred. No.:              7.66e-264     Length:        2930
Score:                  4612.00       Matches:       873
Percent Similarity:     94.38%        Conservative:  1
Best Local Similarity:  94.28%        Mismatches:    6
Query Match:            34.90%        Indels:        46
DB:                     18            Gaps:          1

US-09-522-753-5 (1-2517) x US-10-174-014-13 (1-2930)

Qy        1592 IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyr 1611
               :::         ||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 CTTGGCTCACCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTAT 182

Qy        1612 GluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 GAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCC 242

Qy        1632 PheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeu 1651
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 TTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTG 302

Qy        1652 ProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArg 1671
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 CCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGC 362

Qy        1672 GlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThr 1691
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         363 GGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACC 422

Qy        1692 SerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArg 1711
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         423 TCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGG 482

Qy        1712 GlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIle 1731
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         483 GGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATC 542

Qy        1732 IleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrPro 1751
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         543 ATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCA 602

Qy        1752 AlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg 1771
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         603 GCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGC 662

Qy        1772 HisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThr 1791
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         663 CACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACG 722

Qy        1792 SerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGlu 1811
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         723 TCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAA 782

Qy        1812 LysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThr 1831
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         783 AAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACA 842

Qy        1832 GluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgPro 1851
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         843 GAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCC 902

Qy        1852 AlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeu 1871
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         903 GCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTC 962

Qy        1872 GlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGlu 1891
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         963 CAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAG 1022

Qy        1892 ProSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAla 1911
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1023 CCCAGCACGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCC 1082

Qy        1912 ThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThr 1931
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1083 ACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACC 1142

Qy        1932 LeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgPro 1951
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1143 CTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCC 1202

Qy        1952 ArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluPro 1971
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1203 CGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCC 1262

Qy        1972 AlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHis 1991
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1263 GCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCAC 1322
```

```
Qy       1992 AlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspPro 2011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 GCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCG 1382

Qy       2012 ProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPhe 2031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 CCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTT 1442

Qy       2032 SerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGlu 2051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 TCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAA 1502

Qy       2052 GlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuPro 2071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 GGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCC 1562

Qy       2072 LysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnPro 2091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1563 AAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCA 1622

Qy       2092 GlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGlu 2111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 GGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAG 1682

Qy       2112 SerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArg 2131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1683 AGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGG 1742

Qy       2132 ValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHis 2151
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1743 GTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCAC 1802

Qy       2152 ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCys 2171
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1803 CCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGC 1862

Qy       2172 ProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGly 2191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1863 CCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGT 1922

Qy       2192 AlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLys 2211
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1923 GCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAG 1982

Qy       2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1983 ACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATG 2042

Qy       2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2043 ACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAG 2102

Qy       2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2103 ACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTC 2162

Qy       2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2163 TTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAAC 2222

Qy       2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 AAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACG 2282

Qy       2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2283 GAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCG 2342

Qy       2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2343 GTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATG 2402

Qy       2352 GlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeu 2371

Db       2402 ------------------------------------------------------------ 2402

Qy       2372 AsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391

Db       2402 ------------------------------------------------------------ 2402

Qy       2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
                                ||||||||||||||||||||||||||||||||||||||||||
Db       2403 ------------------GGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGC 2444

Qy       2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2445 CGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCC 2504

Qy       2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2505 TCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAG 2564

Qy       2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2565 GACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTG 2624

Qy       2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2625 CAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGCCCCCTC 2684

Qy       2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2685 GCTGGCGCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAG 2744

Qy       2512 ThrLeuSerAspSerGlu 2517
              ||||||||||||||||||
Db       2745 ACACTCTCCGACAGCGAG 2762


Search completed: November 12, 2005, 05:30:12
Job time : 6402 secs
```

```
                        GenCore version 5.1.6
               Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:          November 11, 2005, 09:36:02 ; Search time 1201 Seconds
                                          (without alignments)
                                          3429.233 Million cell updates/sec

Title:           US-09-522-753-5
Perfect score:   13215
Sequence:        1 MSGSTQLVAQTWRATEPRYP..........WDEEPKPLLCSQYETLSDSE 2517

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters:      2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09522753/runat_02112005_173641_20819/app_query.fasta_1.2695
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09522753_@CGN_1_1_186_@runat_02112005_173641_20819 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
-----------------------------------------------------------------------------------
    1   12922   97.8    9053  4  US-09-976-594-306          Sequence 306, App
    2  4165.5   31.5    7912  4  US-09-949-016-3804         Sequence 3804, Ap
c   3   726.5    5.5   33529  3  US-09-144-085-3            Sequence 3, Appli
    4     718    5.4    1922  2  US-08-372-652-9            Sequence 9, Appli
    5     718    5.4    1922  5  PCT-US95-16311-9           Sequence 9, Appli
c   6     698    5.3   14462  4  US-09-902-540-1090         Sequence 1090, Ap
c   7     698    5.3   14462  4  US-09-902-540-9597         Sequence 9597, Ap
c   8     689    5.2 4411529   3  US-09-103-840A-1           Sequence 1, Appli
c   9     685    5.2 4403765   3  US-09-103-840A-2           Sequence 2, Appli
c  10     670    5.1   44377  2  US-08-804-227C-7           Sequence 7, Appli
c  11     670    5.1   44377  2  US-08-804-198-1            Sequence 1, Appli
   12   668.5    5.1 4403765   3  US-09-103-840A-2           Sequence 2, Appli
   13     665    5.0 4411529  3  US-09-103-840A-1           Sequence 1, Appli
c  14   661.5    5.0   43280  2  US-08-804-227C-1           Sequence 1, Appli
c  15     652    4.9   13987  2  US-08-804-227C-13          Sequence 13, Appl
c  16   648.5    4.9   50937  3  US-09-428-517-1            Sequence 1, Appli
   17     641    4.9   28958  1  US-08-258-261B-6           Sequence 6, Appli
   18     641    4.9   28958  1  US-08-456-837-6            Sequence 6, Appli
   19     641    4.9   28958  1  US-08-457-342-6            Sequence 6, Appli
   20     641    4.9   28958  1  US-08-457-646A-6           Sequence 6, Appli
   21     641    4.9   28958  1  US-08-458-076A-6           Sequence 6, Appli
   22     641    4.9   28958  1  US-08-764-233A-4           Sequence 4, Appli
   23     641    4.9   28958  1  US-08-457-335A-6           Sequence 6, Appli
   24     641    4.9   28958  1  US-08-729-214-6            Sequence 6, Appli
   25     641    4.9   28958  3  US-09-028-934-6            Sequence 6, Appli
   26     641    4.9   49377  1  US-08-764-233A-1           Sequence 1, Appli
c  27   631.5    4.8   80161  3  US-09-036-987A-1           Sequence 1, Appli
c  28   631.5    4.8   80161  3  US-09-370-700-1            Sequence 1, Appli
c  29   631.5    4.8   80161  4  US-09-603-207-1            Sequence 1, Appli
c  30   624.5    4.7   68750  3  US-09-335-409-1            Sequence 1, Appli
c  31   624.5    4.7   68750  3  US-09-568-102-1            Sequence 1, Appli
c  32   624.5    4.7   68750  3  US-09-567-969-1            Sequence 1, Appli
c  33   624.5    4.7   68750  3  US-09-568-480-1            Sequence 1, Appli
c  34   624.5    4.7   68750  3  US-09-568-486-1            Sequence 1, Appli
c  35   624.5    4.7   68750  3  US-09-568-472-1            Sequence 1, Appli
c  36   624.5    4.7   68750  3  US-09-567-899-1            Sequence 1, Appli
c  37   622.5    4.7   24459  4  US-09-902-540-5004         Sequence 5004, Ap
   38   622.5    4.7   32241  4  US-09-902-540-1247         Sequence 1247, Ap
c  39   622.5    4.7   71989  3  US-09-443-501A-2           Sequence 2, Appli
c  40     619    4.7   47981  4  US-09-679-279-1            Sequence 1, Appli
c  41   616.5    4.7    8438  1  US-07-945-283-1            Sequence 1, Appli
   42   614.5    4.7   15872  3  US-09-105-537-1            Sequence 1, Appli
   43   614.5    4.7   15872  4  US-09-091-609-1            Sequence 1, Appli
   44   614.5    4.7   15872  4  US-09-091-609-3            Sequence 3, Appli
   45   609.5    4.6   17622  4  US-09-902-540-1125         Sequence 1125, Ap


                                ALIGNMENTS


RESULT 1
US-09-976-594-306
; Sequence 306, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
;  APPLICANT: Furness, Michael
;  APPLICANT:  Buchbinder, Jenny
;  TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
;  FILE REFERENCE: PA-0041 US
;  CURRENT APPLICATION NUMBER: US/09/976,594
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,409
;  PRIOR FILING DATE: 2000-10-12
;  NUMBER OF SEQ ID NOS: 1143
;  SOFTWARE: PERL Program
; SEQ ID NO 306
;   LENGTH: 9053
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6673549 898877.6
;   NAME/KEY: unsure
;   LOCATION: 2006, 2012
;   OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-306

Alignment Scores:
Pred. No.:               0            Length:          9053
Score:                   12922.00      Matches:         2487
Percent Similarity:      98.26%       Conservative:    2
Best Local Similarity:   98.18%       Mismatches:      10
Query Match:             97.78%       Indels:          35
DB:                      4            Gaps:            5

US-09-522-753-5 (1-2517) x US-09-976-594-306 (1-9053)
```

```
Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        520 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 579

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 639

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 699

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 759

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 819

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 879

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 939

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 999

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 1059

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 1119

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 1179

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 1239

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 1299

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 1359

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       1360 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1419

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1420 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1479

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1480 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1539

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1599

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1600 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1659

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||   :::||||||||||||||||||||||||||||||||||||||||||
Db       1660 GAGCAGGAGAACCGGAAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1719

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1720 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1779

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1780 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1839

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1840 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1899

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1959

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGln---GlnGlnGlnGln 499
              |||||||||||||||||||||||||||||||||||||||||||||   |||   ||||||
Db       1960 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCNACAACNGCAGCAG 2019

Qy        500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAG 2079

Qy        520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AAAGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAG 2139

Qy        540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 AACGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGAC 2199

Qy        560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 GAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAA 2259

Qy        580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 GGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAG 2319

Qy        600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2320 CAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAA 2379

Qy        620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2380 ATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGG 2439

Qy        640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2440 ATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2499

Qy        660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2500 CAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCG 2559

Qy        680 ArgArgLysLysLysLysAlaProAlaAlaAla-SerGluGluAlaAlaPheProProVa 699
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db       2560 CGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCCAGCGAGGAGGCTGCATTCCCGCCCGT 2619

Qy        699 lValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGl 719
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2620 GGTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGA 2679

Qy        719 uGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPr 739
```

```
                ||||||||||
Db         2680 GGAGGCTGAA-------------------------------------------------- 2689

Qy          739 oAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAl 759
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2690 -GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGC 2748

Qy          759 aLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAsp-GlyProP 779
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         2749 CAAGGACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGGCCAC 2808

Qy          779 roProGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrPro 798
                ||||||||||||||||||||||| |||||||||||||||   |||    |||||||||||
Db         2809 CCCCAGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCG 2867

Qy          799 AlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaPro 818
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2868 GCCTCTGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCT 2927

Qy          819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGlu 838
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2928 CCTCCTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAG 2987

Qy          839 GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2988 GAGGGGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCC 3047

Qy          859 GluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLys 878
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3048 GAGGAGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAG 3107

Qy          879 AspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGly 898
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3108 GACGCGGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGC 3167

Qy          899 GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3168 GGGAGCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGT 3227

Qy          919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3228 GCTACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTG 3287

Qy          939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3288 TCCCCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAG 3347

Qy          959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         3348 AAGCCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATC---GTC 3404

Qy          979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProPro-Al 998
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3405 ACCAAAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGAC 3464

Qy          998 aProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSe 1018
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3465 CCCACCGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAG 3524

Qy         1018 rProArgGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------- 1033
                ||||||||||||||||||||||||||||||||||||||||||||||
Db         3525 CCCCCGGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGT 3584

Qy         1034 ----------AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpTh 1050
                          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3585 GTTCTTCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGAC 3644

Qy         1050 rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3645 TTCCGGCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCC 3704

Qy         1070 oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         3705 GGACCCCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGA 3764

Qy         1090 pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3765 CACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTC 3824

Qy         1110 rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3825 CTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTC 3884

Qy         1130 rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3885 GGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCAT 3944

Qy         1150 tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3945 GGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGA 4004

Qy         1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4005 GCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCA 4064

Qy         1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4065 GGAGGCGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAA 4124

Qy         1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4125 AGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCAC 4184

Qy         1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4185 CCACGGCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGA 4244

Qy         1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4245 CAGCCCGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTA 4304

Qy         1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4305 CGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTC 4364

Qy         1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4365 CAAGGAGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCAC 4424

Qy         1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4425 CTATGACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCT 4484

Qy         1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4485 CATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCA 4544

Qy         1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4545 CATCCGCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTA 4604

Qy         1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4605 CCTGCGTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTC 4664

Qy         1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4665 ACGGGACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGC 4724

Qy         1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4725 CCATGAGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCG 4784

Qy         1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySe 1450
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4785 CGAGGAGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTC 4844
```

```
Qy       1450 rIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLy 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4845 CATCACGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAA 4904

Qy       1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLe 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4905 GCACGACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCT 4964

Qy       1490 uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSe 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4965 GGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAG 5024

Qy       1510 rArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5025 CCGGCCAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGT 5084

Qy       1530 lProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPh 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5085 GCCTGAGCTGGGCAAGCCGCGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTT 5144

Qy       1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl 1570
              ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db       5145 TGCCGGCCACCTCCCACGAGGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCA 5204

Qy       1570 nGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5205 GGAGGGCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCG 5264

Qy       1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5265 TGAGATCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCC 5324

Qy       1610 oTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLe 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5325 CTATGAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCT 5384

Qy       1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTy 1650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5385 GGCCTTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTA 5444

Qy       1650 rLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIl 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5445 CCTGCCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCAT 5504

Qy       1670 eArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIl 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5505 CCGCGGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACAT 5564

Qy       1690 eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
              |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db       5565 CACCTCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCT 5624

Qy       1710 uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGl 1730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5625 GAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGG 5684

Qy       1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5685 CATCATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCAC 5744

Qy       1750 rProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSe 1770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5745 CCCAGCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAG 5804

Qy       1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh 1790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5805 CCGCCACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCAC 5864

Qy       1790 rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5865 CACGTCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCG 5924

Qy       1810 gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGl 1830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5925 GGAAAAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGG 5984

Qy       1830 yThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerAr 1850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5985 TACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCG 6044

Qy       1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6045 CCCCGCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGC 6104

Qy       1870 aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa 1890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6105 CCTCCAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGT 6164

Qy       1890 lGluProSerLysProThrValLeu---------ArgSerThrSerThrSerSerProVa 1907
              ||||||||||   ||||||||||||         ||||||||||||||||||||||||||
Db       6165 GGAGCCCAGCACGCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCCTCACCCGT 6224

Qy       1907 lArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGl 1927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6225 TCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGG 6284

Qy       1927 yValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaAr 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6285 GGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCG 6344

Qy       1947 gProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSe 1967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6345 GCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTC 6404

Qy       1967 rGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProPr 1987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6405 CGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCC 6464

Qy       1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAl 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6465 TGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGC 6524

Qy       2007 aSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGl 2027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6525 CAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCA 6584

Qy       2027 nSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSe 2047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6585 AAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAG 6644

Qy       2047 rTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAs 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6645 CTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGA 6704

Qy       2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuAr 2087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6705 CAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCG 6764

Qy       2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6765 GCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCG 6824

Qy       2107 gProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLy 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6825 GCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAA 6884

Qy       2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6885 AGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTA 6944

Qy       2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6945 CACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCC 7004

Qy       2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7005 TGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCC 7064

Qy        2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerPr 2207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7065 CCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCC 7124

Qy        2207 oGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerPr 2227
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7125 AGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCC 7184

Qy        2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7185 ACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCG 7244

Qy        2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGl 2267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7245 GGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCA 7304

Qy        2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLy 2287
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7305 GCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAA 7364

Qy        2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7365 GCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAG 7424

Qy        2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7425 CCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTA 7484

Qy        2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7485 TAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAG 7544

Qy        2347 gLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAl 2367
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7545 AAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGC 7604

Qy        2367 aPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAs 2387
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7605 TTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGA 7664

Qy        2387 pGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGl 2407
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7665 CGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGG 7724

Qy        2407 yArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPr 2427
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7725 CAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCC 7784

Qy        2427 oProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAs 2447
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7785 ACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAA 7844

Qy        2447 nArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLe 2467
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7845 CCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCT 7904

Qy        2467 uIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGl 2487
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7905 GATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGG 7964

Qy        2487 ySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCy 2507
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7965 CAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTG 8024

Qy        2507 sSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||
Db        8025 CTCGCAGTACGAGACACTCTCCGACAGCGAG 8055

RESULT 2
US-09-949-016-3804
; Sequence 3804, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3804
;   LENGTH: 7912
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-3804

Alignment Scores:
Pred. No.:                9.68e-183      Length:         7912
Score:                    4165.50        Matches:        1103
Percent Similarity:       53.57%         Conservative:   344
Best Local Similarity:    40.84%         Mismatches:     777
Query Match:              31.52%         Indels:         477
DB:                       4              Gaps:           109

US-09-522-753-5 (1-2517) x US-09-949-016-3804 (1-7912)

Qy          16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
               :::   ||||||||||||||||||:::   |||               |||
Db         289 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 348

Qy          36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
               :::      :::   :::|||:::      |||               ||||||   |||
Db         349 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 408

Qy          54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                 |||      ||||||||||||||||||||||||||||||
Db         409 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 468

Qy          74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
               ||||||::::::|||   ||||||         |||               :::   |||
Db         469 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 522

Qy          94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                              :::   :::   :::||||||||||||||||||||||||   :::
Db         523 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 582

Qy         114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                  |||                                 ||||||:::   |||   |||
Db         583 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 642

Qy         123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                           |||      |||   |||   |||      ||||||   :::
Db         643 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 690

Qy         143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
               ||||||   |||   |||   ||||||   |||                  :::
Db         691 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 750

Qy         162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                     |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db         751 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 810

Qy         182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
               ||||||      |||||||||||||||   ||||||||||||||||||||||||||||||
Db         811 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 870
```

```
Qy          202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
                |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db          871 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 930

Qy          222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
                ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db          931 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAAAAAGCAGAAGAAGCT 990

Qy          242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
                |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db          991 CATAAAATTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 1050

Qy          262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIle 281
                ||||||:::   ||||||||||||||||||   ||||||   ||||||||||||||||||
Db         1051 GATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATT 1110

Qy          282 LeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArg 301
                |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db         1111 TTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAGCGT 1170

Qy          302 TyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArg 321
                |||||||||||||||||||||   ||||||||||||:::|||||||||||||||||||||
Db         1171 TATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGG 1230

Qy          322 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArg 341
                |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         1231 AGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGA 1290

Qy          342 LysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeu 361
                ||||||||||||   |||||||||   |||   ||||||||||||||||||:::||||||
Db         1291 AAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGGCAGAGGGGAGCTGGTCTT 1347

Qy          362 SerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGlu 381
                |||   :::   ||||||||||||||||||:::|||||||||||||||||||||||||||
Db         1348 TCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGGCTCTCTGAG 1407

Qy          382 GlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAsp 401
                |||||||||   |||||||||||||||||||||:::|||||||||||||||::::::|||
Db         1408 CAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGATTCCACCTATGATGTTTGAT 1467

Qy          402 AlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLys 421
                |||:::|||:::|||:::|||||||||||||||||||||||||||   ||||||||||||
Db         1468 GCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCTTATGGAGGACCCTATGAAA 1527

Qy          422 ValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArg 441
                ||||||||||||||||||   ||||||:::|||::::::   |||||||||   |||:::
Db         1528 GTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAG 1587

Qy          442 GluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLys 461
                :::||||||:::|||||||||||||||||||||||||||||||||:::||||||||||||
Db         1588 GACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCATCATACTTGGAGAGGAAG 1647

Qy          462 ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSer 481
                :::|||   :::|||||||||||||||||||||||||||||||||||||||||||||:::
Db         1648 AGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAAAATGAGAATTATAAAGCC 1707

Qy          482 LeuValArgArgSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||:::|||   :::|||||||||::::::||||||      :::
Db         1708 CTCGTCAGAAGGAATTATGGAAACGCAGAGGCAGAAACCAGCAAATTGCTCGACCCTCG 1767

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                |||:::::::::   :::::::::::::::         ::::::   :::   ||||||
Db         1768 CAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAAGCAGAAAAAACAGAAAAA 1821

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                   |||:::||||||:::   |||   ||||||:::|||::::::         ||||||
Db         1822 AAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAAGAAGACTCCAAAGAAAAT 1881

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                   ||||||         |||:::|||   |||   |||:::   |||:::   :::|||
Db         1882 ACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA---GAAGAAACTGAGGAA 1929

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                :::|||            :::|||||||||||||||||||||||||||||||||||||||
Db         1930 AGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGTCAGGGCCGCCGTAAGGGC 1989

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||   |||||||||   :::      |||
Db         1990 CGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGTGCTGCAGCCGCAGCGGCT 2049

Qy          601 Ser------------------------------AlaGluLeuAlaSerMetGluLeuAsn 610
                :::                                 |||      |||   |||
Db         2050 ACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCCATTTCTACAGAGCCTGTG 2109

Qy          611 GluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHis 630
                |||:::|||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db         2110 GAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGTCTAGTAGAACAT 2169

Qy          631 GlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLys 650
                ||||||||||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||
Db         2170 GGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAAAGTGAAGCTCAATGTAAA 2229

Qy          651 AsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLys 670
                |||||||||||||||||||||:::|||   |||||||||   :::|||||||||||||||
Db         2230 AACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAACCTCTTACAGCAGCATAAA 2289

Qy          671 LeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAla 690
                   |||      ::::::            |||         :::      :::   |||
Db         2290 CAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCTCAATGTGAAAGTGTCGCT 2349

Qy          691 SerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyVal 710
                |||      :::|||               |||||||||::::::|||||||||
Db         2350 TCCACTGTTTCTGCT------------CAGGAGGATGAAGATATTGAAGCCTCC------ 2391

Qy          711 SerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGlu 730
                      ||||||||||||      |||::::::|||
Db         2392 ------AATGAAGAAGAAAATCCAGAAGACAGCGAA------------------------ 2421

Qy          731 ValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSer 750

Db         2421 ------------------------------------------------------------ 2421

Qy          751 IleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProPro 770
                                  ||||||   |||
Db         2422 ---------------GTTGAAGCTGTCAAG------------------------------ 2436

Qy          771 AlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThrSer 790
                                           |||         :::|||         ||||||
Db         2437 ---------------------------CCCAGCGAGGACAGTCCTGAAAATGCTACTTCT 2469

Qy          791 ArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProPro 810
                |||         ||||||            |||            ||||||
Db         2470 CGAGGAAACACAGAACCTGCGGTTGAGCTTGAG------------CCCACCACGGAAACT 2517

Qy          811 AlaPro---ProSerProSer---AlaProProProValValProLysGluGluLysGlu 828
                ||||||      |||||||||   |||   |||         |||   |||::::::
Db         2518 GCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACAAAACCAGCTGAAGATGAAAGT 2577

Qy          829 GluGluThrAlaAlaAlaProProVal--------------------------------- 837
                   ||||||               :::
Db         2578 GTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACAGCAGAGCAGATGGATGTAGAT 2637

Qy          838 ------------------GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeu 851
                                  |||||||||            ||||||   |||   :::
Db         2638 CAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGATCCCCCACCCGCTACCAAA--- 2694

Qy          852 AlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAlaGlu 871
                      |||:::         |||      |||                  :::   |||
Db         2695 ---GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAAAACCATGCATCTAAAGTTGAA 2751

Qy          872 GluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeu 891
                            |||   :::|||   :::   |||
Db         2752 GGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC------------------------ 2787

Qy          892 LysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGlyAla 911
                   :::||||||   |||                        |||:::         |||
Db    2788 ---AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTGGTAGCTCAGCAAATAAATGCC 2844

Qy     912 ------------ProGln---AspSerAspSerSerAlaThrCysSerAlaAspGluVal 926
                       ||||||   |||:::||||||||||||||||||||||||||||||
Db    2845 CAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCCACGTGCAGCGCTGATGAG--- 2901

Qy     927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro---------ArgProSer 943
              :::   :::|||      :::   :::   :::   |||         :::||||||
Db    2902 ---GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTTCCTATGGACTCAAAGCCTTCA 2958

Qy     944 LeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------ProLeu 961
           ||||||   |||||||||               :::||||||   |||      ||||||
Db    2959 CTGTTAAACCCCACTGGATCTATACTCGTC---TCATCTCCGTTAAAACCAAATCCACTG 3015

Qy     962 AspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIle--------------- 976
           ||||||   ||||||:::   |||||||||   |||||||||:::
Db    3016 GATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCACCAATGGTATCCTGCACCCCA 3075

Qy     977 ---------------------------------------------------GlnValThr 979
                                                                 :::
Db    3076 TGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCTCTCTACCAGCGACACATTAAA 3135

Qy     980 LysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaPro 999
              :::||||||            :::
Db    3136 GCAATGCATGAGTCAGCACTCCTGGAG--------------------------------- 3162

Qy    1000 ProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerPro 1019
                    :::      |||   :::   |||         :::      :::||||||
Db    3163 ---GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAATGTAGAAGTTCTACAAGTCCA 3219

Qy    1020 ArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAla 1039
              |||   |||:::||||||            ::::::|||
Db    3220 TGTGGCACATCCAAGAGTCCA------------AACAGAGAG------------------ 3249

Qy    1040 GlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPro 1059
                                      |||
Db    3250 ---------------------------TGG------------------------------ 3252

Qy    1060 ArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPro 1079
              ||||||::::::   :::||||||                     :::      |||
Db    3253 ---GAAGTCCTTCAGCCTGCTCCACAT------------CAAGTGATAACTAATCTCCCT 3297

Qy    1080 ProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPro 1099
              |||      ||||||               |||   ||||||      |||   |||
Db    3298 GAAGGCGTTCGGCTTCCG---------------ACAACTCGACCAACCAGGCCACCGCCC 3342

Qy    1100 ProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGlu 1119
           |||   |||   :::            ::::::|||   |||   ||||||   :::
Db    3343 CCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAAAAA---CCATCTTTATA--- 3396

Qy    1120 ArgGlnIleGlyAlaIleSerGlnGly---------------MetSerValGlnLeuHis 1134
                    |||:::||||||||||||               ::::::
Db    3397 ---ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACTTATTTGACTTCTCATAATCAG 3453

Qy    1135 ValProTyrSerGluHisAlaLysAlaPro---ValGlyProValThrMetGlyLeuPro 1153
                 |||::::::            |||   ||||||   :::::::::|||||||||
Db    3454 GCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTGGGATCTATCTCTCTTGGACTGCCA 3513

Qy    1154 LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSer 1173
                    :::      |||   |||            :::|||||||||:::   |||
Db    3514 CGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTACATCAAGCAGGAAGAATTTTCT 3573

Qy    1174 ProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSer 1193
           ||||||   |||         ||||||   |||   |||      ||||||
Db    3574 CCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC---AGGGCCCAACATGAAGGT 3630

Qy    1194 ValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIlePro 1213
           |||:::||||||||||||   |||::::::      ||||||||||||:::|||   |||
Db    3631 GTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGAAGTATAACTCGGGGAACTCCA 3687

Qy    1214 SerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThr 1233
           ::::::::::::      :::|||      :::   |||||||||||||||   ||||||
Db    3688 ACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGGGGCTCTATCACTCAGGGCACC 3747

Qy    1234 ProAla------------------------AspValLeuTyrLysGlyThrIleThrArg 1245
           ||||||                        :::   |||   ||||||:::|||:::|||
Db    3748 CCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTGGTGAAGGGGTCCATTTCGAGA 3807

Qy    1246 IleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLys 1265
           :::      |||||||||            ::::::|||||||||::::::      |||
Db    3808 ATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGCAGAGAGGAAGCTGCATCCAAA 3861

Qy    1266 GlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSer 1285
           ||||||||||||||||||||||||   ||||||:::|||||||||:::         :::
Db    3862 GGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTGTCATATGATAATATTAAGAAT 3921

Qy    1286 ValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAla 1305
                                :::|||      |||            ||||||   :::
Db    3922 GCC---------------CGAGAAGGGACTAGGAGTCCAAGAACAGCTCATGAAATCAGT 3966

Qy    1306 AlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer------ 1323
                 ||||||:::|||:::   :::||||||   :::   :::   :::|||
Db    3967 TTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATAAAGCAAGGGATGTCAATGAGG 4023

Qy    1324 ------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArg 1339
                       ||||||   :::|||||||||:::   ||||||:::|||      |||
Db    4024 GAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGCCGAGCATTACCC------AGG 4077

Qy    1340 HisSerProHis---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358
              |||||||||      |||||||||:::      :::   |||||||||   ||||||
Db    4078 GGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTGTCTGGCTCCATAATGCAGGGG 4137

Qy    1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378
              ||||||:::      |||:::   ||||||   |||   :::      |||   :::
Db    4138 ACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT---AAATATCCCAAACAAATT 4194

Qy    1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
           |||||||||                  ||||||   |||            |||
Db    4195 AAAAGGGAAAGT---------------CCTCCCATACGAGCATTTGAAGGTGCCATT--- 4236

Qy    1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
                                |||   ||||||   ::::::|||   :::   |||:::
Db    4237 ------------------ACCAAAGGAAAACCA---TATGATGGC---ATCACCACCATC 4272

Qy    1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu--------------- 1433
           ||||||   |||||||||||||||||||||||||||:::::::::
Db    4273 AAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAAGATATTTTAACTCAGGAAAGT 4332

Qy    1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
           |||   |||||||||:::      :::   ||||||:::   ||||||||||||:::|||
Db    4333 CGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATAATTGAGGGTTCCATTTCCCAG 4392

Qy    1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
           |||||||||:::|||:::|||         |||         :::   ||||||:::|||
Db    4393 GGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAATCTGCCATCAAACACAATGTC 4449

Qy    1474 ArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
           :::|||||||||      |||   :::            :::   ||||||:::::::::
Db    4450 AAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGAATGCCTCCGCTGGAAATTGTG 4509

Qy    1494 AlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu---------------GluSer 1507
              :::   :::   :::||||||         ||||||               |||:::
Db    4510 CCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAGACC 4569

Qy    1508 LeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaPro 1527
           ::::::||||||      :::      |||||||||   |||:::   |||
Db    4570 GTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCCTCCGTTCTTAGGTCCACA--- 4626

Qy    1528 ValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGly 1547
                 :::   |||      |||   :::   ||||||      |||:::|||
Db    4627 ------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCTGGGATTTATGATGACACCAGT 4680

Qy    1548 Ala---------------ProPheAlaGlyHisLeuProArgGlySerProValThrMet 1562
           |||                  :::         :::   ||||||||||||:::
Db    4681 GCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCCAGAGGCTCACCCATG------ 4734
```

```
Qy          1563 ArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAla---SerGln 1581
                                |||      :::   ::::::||||||:::|||:::   :::
Db          4735 ---------ATGAACAGAACTTCTGATGTTACAATTTCTTCTAACAAGTCTACCAATCAT 4785

Qy          1582 AspArgLysLeuThrSerThrProArgGlu------------IleAlaLysSerProHis 1597
                 :::||||||   |||   ||||||   :::               ||||||||||||
Db          4786 GAAAGGAAATCGACACTGACCCCTACCCAGAGGGAAAGTATCCCAGCGAAGTCTCCAGTG 4845

Qy          1598 SerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGly 1617
                       |||               |||      ||||||::::::         ||||||
Db          4846 CCTGGGGTGGACCCTGTCGTGAGCCAC------AGTCCGTTTGATCCCCATCACAGAGGC 4899

Qy          1618 ValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIle 1637
                    :::      ::::::||||||||||||:::|||         ||||||   ::::::
Db          4900 AGCACTGCAGGCGAGGTTTATCGGAGCCACCTGCCCACGCACTTGGATCCA---GCCATG 4956

Qy          1638 ProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHisLeuAla 1656
                 |||            ||||||   |||||||||||||||         |||   |||:::
Db          4957 CCTTTTCACAGGGCTTTGGATCCTGCAGCGGCTGCTTACCTGTTTCAGAGACAGCTTTCA 5016

Qy          1657 ProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThr 1676
                 |||   |||   ||||||      |||      |||
Db          5017 CCAACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC------------------------ 5052

Qy          1677 AlaAlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMet 1695
                    |||:::||||||   ||||||||||||:::|||||||||||||||||||||||||||
Db          5053 ---GCAATGGAGAACACAAGACAGACAATCTTAAATGATTACATTACCTCACAACAGATG 5109

Qy          1696 HisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerPro 1715
                       |||                     |||   |||:::   |||||||||||||||
Db          5110 CAAGTGAACTTG------------------CGTCCAGATGTGGCCAGAGGACTCTCCCCA 5151

Qy          1716 ArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSer 1735
                 ||||||      |||   |||   |||   |||      ||||||||||||||||||:::
Db          5152 AGAGAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGAGGAATCATTGACCTGACC 5208

Qy          1736 GlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMet 1755
                    :::|||         :::|||||||||         ||||||            |||
Db          5209 AATATGCCTCCA---ACAATTTTAGTGCCTCATCCAGGGGGAACAAGCACTCCTCCCATG 5265

Qy          1756 AspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg---HisSerSer 1774
                 ||||||:::   |||:::|||               |||      |||   ::::::|||
Db          5266 GACAGAATCACTTATATTCCTGGTACACAGATTACTTTCCCTCCCAGGCCGTACAACTCT 5325

Qy          1775 SerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSer 1794
                 :::   :::|||||||||   ||||||||||||                  :::|||:::
Db          5326 GCTTCCATGTCTCCAGGACACCCAACACACCTT---------GCAGCTGCTGCAAGTGCT 5376

Qy          1795 GluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIle 1814
                 ||||||||||||:::|||:::||||||::::::|||:::|||||||||         |||
Db          5377 GAGAGGGAACGGGAACGGGAGCGGGAGAAGGAGCGGGAGCGGGAACGG---------ATT 5427

Qy          1815 LeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSer 1834
                       :::::::::   :::                  |||||||||:::||||||
Db          5428 GCTGCAGCTTCCTCCGACCTC------------TACCTGCGGCCAGGCTCAGAACAG--- 5472

Qy          1835 SerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHis 1854
                                                              ||||||   ||||||
Db          5473 ---------------------------------------CCTGGCCGACCTGGCAGTCAT 5493

Qy          1855 SerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGlnGln 1873
                    :::            ||||||   |||   |||||||||:::      |||||||||
Db          5494 GGATATGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAGGAGACCATGTTGCAACAG 5550

Qy          1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
                 ||||||||||||         |||            :::||||||   ::::::|||:::
Db          5551 AGACCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATCACACCTTTGGATCCAACT 5610

Qy          1894 LysProThrValLeuArgSerThrSerThrSerSerProVal------------ArgPro 1909
                             :::                     |||                  |||
Db          5611 GCTCAGCTACGAATCATGCCACTGCCTGCTGGGGGCCCTTCAATAAGCCAAGGCCTGCCA 5670

Qy          1910 AlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyr 1929
                 |||:::   :::      |||            |||         :::|||
Db          5671 GCCTCCCGTTACAACACTGCTGCGGATGCC---CTGGCTGCTCTTGTGGATGCTGCAGCT 5727

Qy          1930 ProThrLeuMetGluProValLeuLeuProLys------------GluAlaProArgVal 1945
                                   |||         |||            ||||||   |||:::
Db          5728 TCTGCACCCCAGATGGATGTGTCCAAAACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTA 5787

Qy          1946 AlaArgProGluArgProArgAlaAsp--------------------------------- 1954
                             |||   |||:::
Db          5788 GAAGAAAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAA 5847

Qy          1955 ---------------------------------------ThrGlyHisAlaPheLeuAla 1961
                                                        |||      ||||||   :::
Db          5848 ACCCTGGAGGTGGAGAAGAGATCTGTTCAGTGTTTATACACTTCTTCAGCCTTTCCAAGT 5907

Qy          1962 LysProProAlaArgSerGlyLeuGluProAlaSerSer---------ProSerLysGly 1978
                       |||               :::|||   ||||||                  |||
Db          5908 GGCAAGCCC---------------CAGCCTCATTCTTCAGTAGTTTATTCTGAGGCTGGG 5952

Qy          1979 SerGluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrPro 1998
                    :::      |||      ||||||   |||   :::            ||||||
Db          5953 AAAGATAAAGGGCCT------CCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCAGA 6006

Qy          1999 AlaLysAsn------LeuAlaProHisHisAlaSerProAspProProAlaProProAla 2016
                    |||            |||   :::                              |||
Db          6007 GGGAAGACTACCATTACTGCAGCTAACTTCATAGACGTGATCATCACCCGGCAAATTGCC 6066

Qy          2017 SerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeu 2036
                 |||      |||               :::||||||         |||
Db          6067 TCGGACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT------------ 6114

Qy          2037 GluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProVal 2056
                          ||||||      |||                  :::   :::|||   :::
Db          6115 ------AGTAGCTTATCTTCTCACAGGTATGAAACACCTAGCGATGCTATTGAGGTGATA 6168

Qy          2057 SerProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGlu 2076
                 ||||||   |||||||||:::         :::|||                  |||
Db          6169 AGTCCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAGACCTATCAGCCAGAGGTT 6228

Qy          2077 LeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeu 2096
                 :::      :::      |||   :::         :::      ||||||
Db          6229 GTTAAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATATGAAGGACCA--------- 6279

Qy          2097 GlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSer 2116
                                   |||   |||   ||||||            |||   |||
Db          6280 ------------------TTACATCACTATCGACCA---------CAGCAGGAATCACCA 6312

Qy          2117 SerProLeuLeuGlnThrAlaProGly------------------------ValLysGly 2128
                 ||||||      |||      |||                           |||
Db          6313 TCTCCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGAATGGGGCAAGTGCCCAGG 6372

Qy          2129 HisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThr 2148
                       |||::::::|||||||||   ||||||   ::::::||||||||||||:::
Db          6373 ACCCATCGGCTGATCACACTTGCTGATCACATCTGTCAAATTATCACACAAGATTTTGCT 6432

Qy          2149 ArgHisHis---------------ProGlnGlnLeuSerAlaProLeuProAlaProLeu 2163
                 |||:::                  |||||||||         |||      :::
Db          6433 AGAAATCAAGTTTCCTCGCAGACTCCCCAGCAG---------CCTCCTACTTCTACATTC 6483

Qy          2164 TyrSerPheProGlyAla------SerCysProValLeuAspLeuArgArgProProSer 2181
                    :::   |||   |||      |||   ||||||         |||         |||
Db          6484 CAGAACTCACCTTCTGCTTTGGTATCTACACCTGTG---------AGGACTAAAACATCA 6534

Qy          2182 AspLeuTyrLeuProProProAsp---------------HisGlyAlaPro---AlaArg 2195
                 :::   |||   |||                        |||      |||   :::|||
Db          6535 AACCGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCATCAAAGACCAGGTTCAAGG 6594

Qy          2196 GlySerProHis---------------SerGluGlyGlyLysArgSerProGluProAsn 2210
                    ||||||                  |||   |||   :::            |||
Db          6595 GTCTCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGGCCTGGAAAATCCCCAGAG 6654

Qy          2211 LysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGly 2230
                 ::::::   |||               :::      ||||||:::|||||||||:::
```

```
Db          6655 AGGAGTCACGTC---------TCTTCGGAGCCCTACGAGCCCATCTCCCCACCCCAG--- 6702

Qy          2231 MetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGlu 2250
                                |||   :::               ||||||   :::   |||
Db          6703 ---GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTCTTGTCTCAGAGGGGCGCA 6759

Qy          2251 GlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAla 2270
                 :::            |||      ::::::|||||||||:::   |||      |||:::
Db          6760 GAGCCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGTATAAGCTACTTGCCTTCA 6819

Qy          2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIle 2290
                 ||||||:::||||||   |||:::   |||   |||||||||||||||||||||||||||
Db          6820 TTCTTCACCAAGCTT---GAAAATACATCACCCATGGTTAAATCAAAGAAGCAGGAGATT 6876

Qy          2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly 2310
                    :::|||||||||:::            :::   :::         :::|||||||||
Db          6877 TTTCGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGA 6936

Qy          2311 ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGln 2330
                 |||||||||||||||:::||||||:::|||   :::|||   ::::::   |||
Db          6937 ACTGAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCAGTTAGCTCTAGAGGCCAT 6996

Qy          2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeu 2350
                 :::      :::   ||||||   |||:::|||||||||   ||||||||||||||||||
Db          6997 TCTTTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGACATTATCAGGAAGGCTCTC 7053

Qy          2351 MetGlyLysTyrAspGlnTrpGluGlu------------------SerProProLeuSer 2364
                 ||||||   :::|||         |||                  |||   |||:::
Db          7054 ATGGGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTCATGTCCCAGCCTATGGGA 7113

Qy          2365 AlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThr 2384
                                |||      |||:::   |||:::               :::|||
Db          7114 GTA---------GTGCCTGGTACTGCCAACACCTCAGTT---------------GTGACC 7149

Qy          2385 AlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGly------Lys 2402
                 :::   :::   |||   :::            ||||||      ||||||      |||
Db          7150 AGTGGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCATTCAGGAGGAGTTTGCAAA 7209

Qy          2403 AlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro------Gly 2420
                    |||:::      :::   :::|||||||||:::|||||||||   |||      |||
Db          7210 CCAAAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCTATACCTGGGCAAGGC 7269

Qy          2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
                             :::||||||   ||||||||||||||||||||||||||||||   :::
Db          7270 TACTTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCAGAAGGGGATTACCAT 7329

Qy          2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
                 |||:::||||||               |||||||||||||||||||||   |||||||||
Db          7330 AGGCAGACGCCA------GGGTGGGCCTGGGAAGACAGGCCCTCTTCAACAGGCTCAACT 7383

Qy          2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
                    ||||||||||||||||||   ||||||:::         :::   ::::::||||||
Db          7384 CAGTTTCCTTATAACCCTCTGACTATGCGGATG---------CTCAGCAGTACTCCACCA 7434

Qy          2481 ProProGlyLeuProAlaGlySerGly---ProLeuAlaGlyProHis---------His 2496
                    |||         |||   |||            |||   ||||||
Db          7435 ACACCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCTCACCAACAGAACAGG 7494

Qy          2497 AlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSer 2516
                    |||:::   ||||||   |||||||||   :::||||||||||||||||||||||||
Db          7495 ATCTGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAGACCCTGTCGGATAGT 7554

Qy          2517 Glu 2517
                 :::
Db          7555 GAT 7557

RESULT 3
US-09-144-085-3/c
; Sequence 3, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
;  APPLICANT: Gustafsson, Claes
;  APPLICANT: Betlach, Mary C.
;  APPLICANT: Ashley, Gary
;  APPLICANT: Julien, Bryan
;  APPLICANT: Ziermann, Rainer
;  TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: 30062-20020.20
;  CURRENT APPLICATION NUMBER: US/09/144,085
;  CURRENT FILING DATE: 1998-08-31
;  EARLIER APPLICATION NUMBER: 09/010,809
;  EARLIER FILING DATE: 1998-01-22
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 33529
;   TYPE: DNA
;   ORGANISM: Sorangium cellulosum
US-09-144-085-3

Alignment Scores:
Pred. No.:                9.66e-24      Length:        33529
Score:                    726.50        Matches:       674
Percent Similarity:       32.36%        Conservative:  314
Best Local Similarity:    22.08%        Mismatches:    1199
Query Match:              5.50%         Indels:        873
DB:                       3             Gaps:          137

US-09-522-753-5 (1-2517) x US-09-144-085-3 (1-33529)

Qy               17 ProArgTyrProPro---HisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
                 ||||||   ||||||      ||||||      |||      :::|||
Db         22686 CCGAGGCCTCCACCGACACCATCGCTCCCCCGGCCTCGCACCCTTGCA------TCAGCC 22633

Qy               36 AspValGlyLeuLeuGluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerPro 55
                 |||:::|||                        |||      ||||||
Db         22632 GACCTCGGG------------------------CGCACACTAGCTTCGCTGCGTCCGCAA 22597

Qy               56 GlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGly 75
                 ||||||         |||      :::|||   :::|||||||||         |||
Db         22596 GGCTCAGCACACCCGCCACATGCGCAGCGCTCAGCTCTCCTATCGAGT---GCCCCAGCA 22540

Qy               76 AsnGluArgSerGln------------------------GluLeuHis------------ 83
                    :::|||:::|||                        :::   |||
Db         22539 GCGCAGCGGGCTCAGCCCCCACGACTCCCACTGACGGTACAGCGCCACTTCCAGCGCAAA 22480

Qy               84 ------------------LeuArgPro---GluSerHisSerTyrLeuProGluLeuGly 96
                                   |||||||||   :::      |||:::|||         |||
Db         22479 CAGCCCGGGCTGCGTGTACTCCGTCCGCTCCAGCAACGCTCCTTCCTCGCTGCCCGCGGC 22420

Qy               97 LysSer---------GluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                 :::                        :::|||   :::      ||||||      |||
Db         22419 CGCGAACACCACCTCTCTCAACCCACGGTCGAGATACGCGTCCAGCGCCTCGCACACCTC 22360

Qy              114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                    :::                                 ||||||      |||   |||
Db         22359 GTCGAACGCCGCACGGAACACGGGGTACACTTCGTAAAGCCTCTTCCCCATCCCGAGCCG 22300

Qy              123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                 |||   |||                                       |||
Db         22299 CTGGCTGCCCTG---------------------------------CCCCGTGAACAGCAC 22273

Qy              143 GlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeuGluLeu 162
                    ||||||   |||   |||         |||   |||                  :::
Db         22272 CGCAAGCTT---CCCGCCTCGCTTCGCCGTCCCGCTCAC---------CACCGCCGCATC 22225

Qy              163 ValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGlu 182
                    |||||||||      :::   :::   :::   :::               |||:::
Db         22224 CGGCCGCCCCGACGACAGCGCGCGAAGACCCTCCACAGCTCCCGCAGCGCTCGCCGCAAG 22165

Qy              183 IleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGluGlu 202
                                   ::::::      :::::::::::::::
Db         22164 CACCGACGCCCGCGACTCGAAGTGCGTCCGGTGCAGCGCCGCCGT------------GCG 22117

Qy              203 GluAlaAlaLysProProGluProGlu---------LysProValSerProProProIle 219
```

```
                    :::||||||                  :::|||:::   |||||||||
Db   22116 CACCACGTCCGACCACCCCACCTCCCCGTGCTCTTCCAGCCACTTCGCCCACCGCCCCGC 22057

Qy     220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsn------ArgLysLys 237
                    ||||||   ||||||               :::         |||::::::
Db   22056 CTGCGCATTCACCGCGGCCTCGTCTCGACCCGACAGCACCAGCGGCAATAGCGCAGGCGC 21997

Qy     238 AlaGluAlaAlaHisArgIleLeuGluGlyLeuGly------------------------ 249
                          |||:::         ||||||
Db   21996 CGCCTCGGCCTCGACAGGCTCGCGCCGCGCCTCGGCCGGCGCTTCTTCGAGGATGATATG 21937

Qy     249 ------------------------------------------------------------ 249

Db   21936 CGCGTTCGTCCCGCTGATCCCGAACGACGACACGCCGGCACGACGGACCCGGCCGTTGCG 21877

Qy     250 ---ProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGlu 268
              |||:::      |||            ||||||:::            :::
Db   21876 CCGCCACGGACGCGCCTCTTGCAGCAGCGACAGCCCGCT---------TCCTTCCCACCC 21826

Qy     269 AsnIleLysIleAsn------------GlnAlaMetArgLysLysLeuIleLeuTyrPhe 284
           |||::::::                  |||   :::|||:::   |||:::|||:::
Db   21825 AATGTGCGGGCTCGGCTGCTCCGCGTGCAGCGTCTTCGGCAGCACCTCGTGCTGCATCGA 21766

Qy     285 LysArgArg---AsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAsp 303
           :::         :::|||||||||
Db   21765 CAGCACCATCTTGATCACGCCCGC------------------------------------ 21742

Qy     304 GlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgArg 323
                                                        :::   |||||||||
Db   21741 ---------------------------------------------CACGCCCGCCGCCGC 21727

Qy     324 AlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLys--- 342
              :::   |||:::||||||                           :::|||:::
Db   21726 CTGCGCATGTCCGAGGTTCGA---------------------------CTTCGACGATCC 21694

Qy     343 ------GlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                 ||||||   |||         ::::::|||:::      ::::::   ||||||
Db   21693 CAGGTACAGCGGACGCTCGGGGCTACGCTCCGGTCCAAACACCTCCGCCAGCGCTCCGGC 21634

Qy     361 LeuSer---MetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
           |||      :::|||   ||||||:::                     :::|||   |||
Db   21633 CTCGATCGGGTCTCCAAGGCTCGTGCCCGTACCGTGCGCCTCCACCGCGTCGATGTCCTC 21574

Qy     380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
                 |||                     ||||||               |||:::|||
Db   21573 GGGCGACAGACCACA------------CGACGACAGCGCCTGCCGGATCACCCGCTGCTG 21526

Qy     400 TyrAspAlaAspGlnGlnArg-IleLysPheIleAsnMetAsnGlyLeuMetAlaAspPr 419
                 |||   ::::::|||       ::::::      :::         |||:::||
Db   21525 GGCAGGGCCGTTCGGCGCCGTCAGACCCTGGCTGCGACCGTCCTGGTTCACCGCAGAGCC 21466

Qy     419 oMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluTh 439
           |                  |||                  |||   |||      :::
Db   21465 ACGGATCACCCCCAGCACACGGTCGCCGTCGCGCTGCGCGTCAGA-CAGCCGCTTCAA-- 21409

Qy     439 rPheArgGluLysPheMetGlnHisProLysAsnPheGly---LeuIleAlaSerPheLe 458
                              |||||||||         |||         |||      ||
Db   21408 ----------------CAACAGCATCCCGCAACCCTCGGCCCAGCCCGCGCCGTCCGCCT 21365

Qy     458 uGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAs 478
           |   |||:::   :::|||      |||      :::   |||            |||
Db   21364 GCACCGAGAAGCTCTTGCACCGACCGTCGCGGACATCCCCTT-------------GAGAC 21317

Qy     478 nTyrLysSerLeuValArgArg------SerTyrArgArgArgGlyLysSerGlnGlnGl 496
               :::         :::|||      ::::::   |||            |||:::||
Db   21316 GGCTGAACTCCACGAACAGCGCGGGGTGCTCATCACCGTCACCCCGCCGGCCAGCGCCA 21257

Qy     496 n------------------GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln-- 509
           |                  |||::::::   :::|||   |||:::   |||::::::
Db   21256 GGTCGCATTCGCCCTGGCGCAGCGCCGTGCACGCCAGGTGCAGCGACACCAGCGACGACG 21197

Qy     510 ----------------------------------GlnProMetProArgSerSerGlnGl 518
                                             ||||||      |||
```

```
Db   21196 AGCACGCCGTGTCCACCGTCATCGCTGGGCCCTGCAGCCCCAG---CGCGTAGGCCACCC 21140

Qy     518 uGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluVa 538
               :::::::::      :::      :::|||            :::      :::::
Db   21139 GGCCCGAGAGGATGCTCGCCGCGCTCCCGATGCCCCGGTACCCGTCCAGCTCCTTCAGAT 21080

Qy     538 lGluAsnAspLysGluAspLeuLeuLys---GluLysThrAsp---AspThrSerGlyGl 556
           :::::::   |||      :::::::::      :::|||         :::   |||
Db   21079 CGACAGTATGAAGAGCACCGTAGTCCGAGCCCATCGAACCCAGGTACACTCCGGTGGCGC 21020

Qy     556 uAspAsnAspGluLysGluAlaValAlaSerLys-------------------------- 567
                        :::   |||   |||
Db   21019 TCCCGCTCAGCGCCGACGGGCGCACGCCGGCACGCTCCAATGCCTCCCACGCCGTCTCCA 20960

Qy     568 -------------------------------------GlyArgLysThrAlaAsnSerGl 575
                                                   |||:::   :::
Db   20959 GCACCAGCCGATGCTGCGGATCCATCGACTGCGCCTCGCGCGGCGATA-TCCCGAAGAAG 20901

Qy     575 nGlyArgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAl 595
               ||||||   ||||||                        |||   |||
Db   20900 TTCGCGTCGAACAGGTCGATGTCCCGAACAAACCCACCTTCGCGCGTAGCTCTTGCCT 20841

Qy     595 aIleThrProGlnGlnSer---------------------AlaGluLeuAlaSerMetGl 608
                  |||   |||                        :::   |||      :::
Db   20840 GCGGCCTCCGGATCAGGGTCGTAGACCTCGAACCCGTCCCAGCGCGCTGGGAGGCCTTCG 20781

Qy     608 uLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLe 628
                     |||   |||      :::               |||   :::   |||
Db   20780 ATCGCGTCCTTCCCCTCGGCCAAGAGACGCCAGTAGTCGTCCGGCGTCGCAACGCCTCCC 20721

Qy     628 uGluHisGly-ArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerG 648
                  ||| |||:::               |||:::|||         |||   :::
Db   20720 GGCAGCCGGCACGCCATCGACACGATCGCGATCGGCTCGTCTTTCCCTGCACGACGACGC 20661

Qy     648 lnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeu---------------- 662
                :::      :::   :::   :::   |||:::   |||
Db   20660 ACCCGCGCCCGCGCTTCCTTCACCTGCAGCCCCGAGAACGCCTGCTTCAGCAGCAGCTCT 20601

Qy     663 -----------------------------AspGluIleLeuGlnGlnHisLysLeuLysM 673
                                        |||:::      :::|||   :::|||
Db   20600 GCGATCGCCCGCGGCGTCGGGTAGTCAAAGACCAGCGTCGCCGGCAAGGACGTCTCGGTA 20541

Qy     673 etGluLysGluArgAsnAlaArgArgLysLysLysLys---------------------- 685
                :::|||      |||:::|||   :::   :::
Db   20540 CGGGACGTGAGGCGGTTGCGCAGCGCCACCGCCATCAGCGAGTCGAGCCCCAGCTCCTTG 20481

Qy     686 -----------------------AlaProAlaAlaAlaSerGluGluAlaAlaPheProP 698
                                  :::|||      :::         ||||||      |
Db   20480 GCACCTGATCCGCTGCCACAGCCTCGCCGCGCTGAAGCCCTGCCACGGCCGCAACCTCGC 20421

Qy     698 roValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetV 718
           ||                        :::         ||||||:::   |||
Db   20420 CCCGCACCAGCTCGACGAGCGCATTCAGCCGCTCCGCCTCCGGCAGCGCCGAGAGGCGCT 20361

Qy     718 alGluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerG 738
                   |||||||||                  :::      |||            |
Db   20360 CGCGGAGCGCCGAGGCGTCTCGCCTCGTTGCCGAGGACGCCTTGCGCAAG------CTCG 20307

Qy     738 lyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSer-ProHisThrGlu 757
           ||      :::      ||||||:::   :::      :::   |||    |||
Db   20306 GGCGCAACAGCGCACGAAACAGCGCCGGCAGCTCGCCGCTGGACTCCAGCCCACGCTGCA 20247

Qy     758 AlaAlaLysAspThrGly------------------------------------------ 763
           ||||||:::|||   |||
Db   20246 GCTGCGCGAGATCCAGGTGCACTGGGACCAGGCTCGCTTCAGGGCGTGAGAGCGCGGCGT 20187

Qy     764 ------GlnAsnGlyProLysProProAlaThrLeu------------------------ 773
                 :::   ||||||   |||||||||:::
Db   20186 CCAGCAGAGCGAGGCCCTCTTCCACCGGCATCGGCACGAACCCATTGCGCCTCATGCGCG 20127

Qy     773 ------------------------------------------------------------ 773

Db   20126 ACAGCTCGGCCTCGCCCAGGTGCGCTGTCATCCCCACGCCCGCCTGTGTCCAGAAGCCCC 20067
```

```
Qy       774 ---------------------------------------------------------Gly 774
                                                                      |||
Db     20066 ACGCCAGGCTCGTCGCCGCGAGCCCGCCGCCGCGCCGGTGCGCCGCGAGCGCGTCGAGGA 20007

Qy       775 AlaAspGlyProPro---------ProGlyProPro------------ThrProProArg 787
                      ||||||            |||||||||               ||||||
Db     20006 ACGTGTTCGCCGCCGCGTAGTTGCTCTGGCCCGCCGCGCCAAACGTACCGGCCGCCGACG 19947

Qy       788 ArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThr 807
             ||||||:::|||   |||   :::|||   |||||||||         ||||||
Db     19946 AGAACAGCACGAACGCCGCGAGATCCAGCTCCCGCGTCAGCTCGTGCAGGTGCAGCGCCC 19887

Qy       808 ProProProAlaProProSerPro---------SerAlaProProProValValProLys 824
                ||||||:::   |||:::|||         :::   |||               :::
Db     19886 CGTCCACCTTCGGCGCCAGCACCCGCGAGAGGCGCTCGGCCGTCTGGGCGGTGAGCACGC 19827

Qy       825 GluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGln--- 843
                   :::               |||||||||         :::         :::
Db     19826 CGTCGTCGAGCACGCCGGCCAGGTGCAGCACCGCGCTCAGCGGACGCGCCGCGTCGATGC 19767

Qy       844 LysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLys 863
             :::|||      |||                  |||   :::   :::   |||
Db     19766 CGGCCAGCACACGCGCGACCTCCTCCCGCTTCGACACGTCGCACGCAGCCACCGTCACCG 19707

Qy       864 SerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAla 883
             |||                              |||:::|||      |||      |||
Db     19706 TCTCGGCGCCGAGCTCCTCGAGTGACTGCACGAGCTCGCGGGCCCCGGGCGCCTCCAGCC 19647

Qy       884 GluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAla 903
                                        |||         :::|||   |||   |||
Db     19646 CGC---------------------GACGCGACGTCAGCACAAGGTGCCTCACCCCGT--- 19611

Qy       904 ThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAla 923
                   |||::::::   |||||||||:::      |||   :::
Db     19610 ------GCGCGCGCACCAGGTGCGCCGCGATCGCCTGACCCAGCTCCCCTGTCCCGCCGG 19557

Qy       924 AspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeu---------------Leu 938
                               :::|||      :::      |||               |||
Db     19556 TCACCAGCACCGTGCCCGCAGGGTCCAGCTCGCGGGCTCCTCGGGTGAGCTCTTCCGCTG 19497

Qy       939 SerPro---ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerPro 957
                |||   |||               |||         |||||||||      :::|||
Db     19496 CCGCCTGCGCGCACCAGGCGCGCGGCCAGCGCAGCGCCCCGCGCAGCGCAAGCTCCG 19437

Qy       958 GlnLysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGln 977
                   |||                              |||      |||      |||
Db     19436 GCTCCGCCG---------------------CCGTCGCCAGCGCCCGCGCCAGCAGCCCAG 19398

Qy       978 ValThrLysValHisGluProProArgGluAspAla------------------------ 989
                   :::         |||   |||      |||
Db     19397 CGTCCACGGGCTCGGTCCCCACATCGATCAAGCGCAGCCGGCGCTCGGGGTGCTCGCTGC 19338

Qy       990 AlaProThrLysProAlaProProAla------ProProProProGlnAsnLeuGlnPro 1007
             ||||||   :::   ||||||   |||         |||   |||         :::|||
Db     19337 GCGCCGCGCGAACAAGCCCCCACAGCGGCGCGCGCGCCAGATCCTGGACGGCGTCGTCCG 19278

Qy      1008 GluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAla 1027
                         |||:::            :::|||      :::   |||:::   |||
Db     19277 GCGCCGCGCCCACCGCGCCCCGCGTCACCCACACCAGCTCGACAGCGTCGAGCCGCGGCT 19218

Qy      1028 ProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProPro 1047
             |||   |||         |||      |||               |||         |||
Db     19217 CCGACAGCCACGCTTGCAGCAGCGACAGCG---------------CCTGCCTCGTCGCCT 19173

Qy      1048 CysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerPro 1067
                         |||         |||:::|||         |||:::|||   |||
Db     19172 ---------CGTGGGACAACGACGCCACTCCCGACTGACTCTGTCCTAAAGCTGTCGTGT 19122

Qy      1068 HisAlaProAspProSerAlaPheSerTyrAlaProProGly------------------ 1081
                   |||         |||   |||            |||
Db     19121 CGACCGCCACCCGCGCAGGCGCGCTCGCGCCCTGCTCGAGGCGCGCAACCAATGCTTCGA 19062

Qy      1082 ---------HisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
                         |||   |||   ||||||            ||||||
Db     19061 GATCGGCAATCGCCTCCGCCCTCAGGGCTT------------CGGCCAGCCTGCCTCGTC 19014

Qy      1099 ProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeu 1118
             ||||||            ||||||            ::::::      |||      |||
Db     19013 CCTCCGCTGCACCGACGACCACCA------------GCGAGCCCGCCTCCAGGGGGGCTG 18966

Qy      1119 GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSer 1138
                |||         |||         |||   |||         :::|||   :::
Db     18965 CCGCGA---------GCTCCACAGGCTGGAAGCTCACCTGATACATATGTTGGCCACCAG 18915

Qy      1139 GluHisAlaLysAlaProValGlyProValThrMetGlyLeu------------------ 1152
                      ::::::               :::   |||
Db     18914 CGTGAACGGCGGTCCGCAGCTGCTCGGCCGTCGCCCGGCGCAGATGTAGAGCGCCGATGC 18855

Qy      1153 ---ProLeuProMetAspProLysLysLeuAlaProPheSer------------------ 1165
                |||   |||                  ||||||:::|||
Db     18854 TCGCCACGGCCT------------------GCCCCGTGGAGTCTGCGACTTGCAGCGAAG 18813

Qy      1166 ---------GlyValLys------------------------GlnGluGlnLeuSerPro 1174
                      ||||||:::                           |||      ||||||
Db     18812 CGGTGTCCTGGTGTGCGCCTCCGTCCTGGAGCTCTACCCGGACCCGAAGCTCGCTCGCCC 18753

Qy      1175 ArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerVal 1194
                   :::|||   |||                  ||||||   |||   |||
Db     18752 CCGTCGCGTGCAGCGCCACGT---------------CCGACCAGGCAAACGGCACGAGCA 18708

Qy      1195 LeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSer 1214
                |||                     |||   ||||||      :::      ||||||
Db     18707 CCTCGTTCG------------------CCCCTGGGCTCTCGGAGAAGGCCGCGACCATCG 18666

Qy      1215 ------ThrArgValProSerAspSerAlaIleThr---TyrArgGlySerIleThrHis 1231
                      |||   ||||||   :::      |||   |||      :::::::::
Db     18665 TGTGGAGCGCGGCGTCCATCAAC-GCCGGGTGCACGCCATACGCCTCCGCCCTGTCCTTC 18607

Qy      1232 GlyThrProAlaAspValLeuTyrLysGlyThrIleThrArg------------Ile-Il 1247
                   |||                  |||:::   ||||||            ::: ::
Db     18606 TCGGTCCCC------------------GGCAGCACCACCCGACCGAAGTAGGTTGTGCCT 18565

Qy      1247 eGlyGluAspSerProSerArgLeuAspArg---GlyArgGluAspSerLeuProLysGl 1266
             :|||      :::|||      |||:::      ||||||
Db     18564 TGGCGCGACAGCTCCACGAGCCCTTGAAACGCCGGGCCGT-------------------- 18525

Qy      1266 yHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerVa 1286
                             |||   :::   |||      :::
Db     18524 -------------AGTGGAGCCCGCGCTCACGCAGCCGCT-------------------- 18498

Qy      1286 lThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro------------ProHi 1302
                                      ::::::   :::||||||            |||
Db     18497 ----------------------CGTAAAGCCCGGACAGGTCCACCTCCTCCGCGCCCGGC 18460

Qy      1302 sGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIl 1322
              :::      :::      |||      |||         |||   :::   |||
Db     18459 ACAGGCCACGTCGACAGCTCGTCGAGCCGACAGGGATGCCGGGGGGCTCGTCCGTCAACA 18400

Qy      1322 eSer------------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIlePr 1336
              |||                           :::   ||||||:::      |||:::
Db     18399 CTCCCGTCGCGTGCTGCACCCACGGCGCGTCCTCCGGGGCCTGCTCAGGCTGGCTGTACA 18340

Qy      1336 oProGlu---------------------ArgHisSerProHisHisLeuLysGlu----- 1347
              |||                        |||   |||   |||||||||:::
Db     18339 GCCCCACCTCGCGCCGGCCCGCCGCGTCCGGCGCTCCGATCATCACCTGCAGCCGCGCCG 18280

Qy      1348 ----------GlnHisHisIleArgGlySer---------------------IleThrGl 1357
                       ||||||         |||
Db     18279 CGCCCTCGGCCAGCACCAGCGGCTCGGCCAGCGTGAGCTCCGACAGCGACCGGCTGCCCA 18220

Qy      1357 nGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGlu------Al 1375
                    ||||||         :::   |||            |||:::         ||
Db     18219 CCGCGCGCCCCGCCGCCAGCACCAGCTCCAGCGTCCCCGTGCCCGGAAGCACCACCTGGC 18160

Qy      1375 aLysLeuLeuLys---------------------ArgGluGlyThrProPro-------- 1385
```

```
             ||||   |||                        |||:::      ||||||
Db     18159 CAAACACCTCATGATCCCGAAGCCACGGATGCTCCGCCAGCGACAGCCGCCCTGTGAACA 18100

Qy      1386 -Pro-ProProProSerArgAspLeuThrGluAlaTyrLysThr---------GlnAlaL 1402
                  |||   ||||||      |||      |||      :::            ||||
Db     18099 GGTGCCCCTCGCCGTCGGCGAGCCTCGTCGCGGCGCCGAGCAGCGGGTGTGCTGCTGCCT 18040

Qy      1402 euGlyProLeuLysLeu-Lys---------------------ProAlaHisGluGlyLeu 1414
             ||   |||   :::||| :::                     ||||||
Db     18039 TGAGCCCCGCCGAGCTCACGTCGCTACGCGCCTTCGAGATATCCAGCCAGTAACGCTGCC 17980

Qy      1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
                   |||         |||                     ||||||   :::   |||
Db     17979 GCTGAAACGCGTACGTCGGCAGCT------------------CCACGACACGAC---CGC 17941

Qy      1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
                   |||      |||::::::      |||
Db     17940 CGTGGCCCGACAGCACCCGTGTCCAGTCCACCT------------------------CGT 17905

Qy      1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
                |||   :::            ||||||   |||   :::      |||   :::|||
Db     17904 GCCCCTGCACGTGCAGCTGCCCCAGCGTCCTGTACAGCTGCGACAACCCAC---CTTCGT 17848

Qy      1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
                      |||||||||            ||||||      |||            :::
Db     17847 CGCGCTGCAGGCTCCCCACCACCA------CCCCCTGCGCCTCCCCGCACGCCGTCGTCA 17794

Qy      1495 AspAlaArgAlaLeuGluArgAla---CysTyrGluGluSerLeuLysSerArgProGly 1513
                |||               |||               :::   ::::::||||||
Db     17793 GCGGCATCGCCAGCACCGGGTGCGCGCTCACCTCCACGAACACACCGTGCCCGTCCTCCA 17734

Qy      1514 ThrAlaSerSerSerGlyGlySerIleAla-Arg------------GlyAlaProValIl 1529
                ||||||::::::   |||      ||| |||               |||||||||::
Db     17733 GCAGCTTCGACAGCGCTCGGTCCAGGCGCACCGTCCGCCGAAGGTTGCGGCACCAGTACT 17674

Qy      1529 eValProGluLeuGlyLysProArgGln-SerProLeuThrTyrGluAspHisGly---- 1547
             :      :::         ||||||    ||||||:::|||   |||   :::|||
Db     17673 CGCCGTCCAGCGCCTCGCCCCGCGACACCTCTCCTGTCACCGTCGAGTAAAACGGCAGCT 17614

Qy      1548 --------------------------------AlaProPheAlaGlyHis---------- 1553
                                             ||||||   :::|||
Db     17613 GCGTCGCCTTCGGCCTGAGCGACGACAGCTTCGCTCCTAGCTCGGGCAGCAGCGCATCCA 17554

Qy      1554 -----------LeuProArgGlySerProValThrMetArgGluProThrProArgLeuG 1570
                                    |||      |||   |||:::   ||||||
Db     17553 TGTGCGCGCTGTGCGACGCGTAGTCGACGTTCACCTTCCGGCAGAACACACCTTCTGCCG 17494

Qy      1570 lnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProA 1590
                     ::::::|||   |||   ||||||               |||         |
Db     17493 TCAGCTCCACCATCAGCCCGTCCACCGCCTCCACGTCACCCGACACCACCGT-CGAGCTC 17435

Qy      1590 rgGluIleAlaLysSer---------------------------------------------- 1595
             ||   :::   :::
Db     17434 GACGTGTTCACCGCCGCTATCGCAAGCGCCTCCCCGTACGGCGCGATGCGCTCCTGCACC 17375

Qy      1596 --ProHisSerThrValProGluHisHis------------------------ProHisProI 1608
                  |||      :::   :::||||||                     ||||||   :
Db     17374 TCCGACACCGGCCGCTCGACCAGCATCATCGCCCCCATCCCCGACCGCTGCCGCACCGCC 17315

Qy      1608 leSerProTyrGluHisLeuLeuArgGlyVal------------SerGlyValAspLeuT 1624
             :::::         |||      |||   :::
Db     17314 TGGCTGCGCAGCGCCACTACCCGCGCTCCCTCCGCAAGCGACAGCGCTCCGCACACCACC 17255

Qy      1624 yrArgSerHisIleProLeuAlaPheAspPro---------------------------T 1635
               |||   |||:::   ||||||      |||                           |
Db     17254 GCCGCCGACACCTCGCCCTGGCTGT-GGCCCACCACCGCCGCAGGCTCCAGCCCCAGCGA 17196

Qy      1635 hrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisL 1655
             ||:::   ||||||      |||      :::      |||         ||||||
Db     17195 CCGCCACGCCGCGGCCAGACCCACGCACATCGCGAACAGCGCGGGCTGCACCACGTCCAC 17136

Qy      1655 euAlaProAsnProThrTyrPro------HisLeuTyrProProTyrLeuIleArgGlyT 1673
               ||||||            |||      ||||||                  :::
```

```
Db     17135 CCGCTCCAGCGACGGCTGCTCCTCCCCGCCATCTCCGCGCAGCACCGACAGCACAGACCA 17076

Qy      1673 yrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerG 1693
               |||   |||||||||      :::|||
Db     17075 GCCCGTCCACGGCCGCAGCGCCTCATCGCACGCCTGCA---------------------- 17038

Qy      1693 lnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyL 1713
                                                :::   |||            |||
Db     17037 -----------------CCGCCTCCGCGAACGCCGCGCTCTGCTCCAGCAGCGCCCGGCC 16995

Qy      1713 euSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleA 1733
               ||||||         ||||||                  |||   |||
Db     16994 CATCCCCGGCCACTGGCTCCCTT------------GGCCGGGGAACACGAACACCACCTT 16947

Qy      1733 spLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaT 1753
               |||:::      |||:::   |||         ||||||            |||||||
Db     16946 GCCTCGCGCACGCGCCGTACCCGCCG------ACACCGCCCGGTGGCCGCGACCCTGCGA 16893

Qy      1753 hrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisS 1773
             |||||                     |||   ||||||   |||      :::|||
Db     16892 CAGCG---------------------CCCGCAGCGCCTCCTCCGCCTCGGACACGCTCGC 16854

Qy      1773 erSerSerProLeu------------------SerProGlyGlyProThrHisLeuThrL 1787
                  :::|||:::                  |||         |||            :
Db     16853 CGCAAGCACCGATGCGCGTGAGGCGAAGTGCGTCCGGTGCAGCGCCGCTGTCCTCACCAC 16794

Qy      1787 ysProThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgA 1807
             ::|||||||||
Db     16793 GTCCGACCACTGCACCTCCCCGTGCTCTCCGAGCCACTTCGCCCAGCGCTCCGCCTGCGC 16734

Qy      1807 spArgGluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleT 1827
                     |||   :::         |||:::|||         :::|||   |||
Db     16733 CCCCACCGAGGCCTCGTCTCGCCCCGACAGCAACAGCGGCATCGCAACACCCTCCGTCGC 16674

Qy      1827 rpArgProGlyThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGly- 1846
                                 |||:::||||||      :::|||   |||
Db     16673 C-AACGGCTCGCGCATTGCCTCCACCGGCTCGCGCGCCGCCTCCACCGGCGCCTCTTCGA 16615

Qy      1847 --------SerSerSerArgProAlaSerHisSerHisAlaHisGlnHisSerPro---- 1862
                        :::   |||||||||               |||      :::|||
Db     16614 GGATGACATGCGCGTT-CGTCCCGCTGATCCCGAACGACGACACGCCCGCGCGCCGCGCC 16556

Qy      1863 --------IleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeu-His 1879
                           |||||||||:::         |||::::::|||:::   ||| |||
Db     16555 CGGCCGTTGCGCCGCCACGGACGCGCCTCTTGCAGCAACGACAGCCCGCTCCCCTCCCAC 16496

Qy      1880 ---AsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeu 1898
                      |||:::               :::|||      |||            :::
Db     16495 CCAATGTGCGGGCTCGGCTGCTCCGCGTGCAGCGTCCGCGGCAGCACCTCGTGCTGCATC 16436

Qy      1899 ArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHis 1918
                ||||||      |||:::|||      |||||||||         ||||||
Db     16435 GACAGCACCATCTTGAGCACACCGGCGACGCCCGCCGCA------GGCCCCGCG------ 16388

Qy      1919 CysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeu 1938
             ||||||   |||   |||                              |||
Db     16387 TGTCCA---GGTTCGACTTCGACGAGCCAAGGTACAGCGGCCGCTCGGCCTTGCGCCCAG 16331

Qy      1939 ProLysGluAlaProArgValAlaArgProGluArg------------------------ 1950
                :::      ||||||      ||||||
Db     16330 GCCCGAACACCTCCACGAGCGCTCCGGCCTCGATCGGGTCGCCCAGGTTCGTGCCCGTCC 16271

Qy      1951 ------------------------ProArgAlaAspThrGly------HisAlaPheLeu 1960
                                     |||||||||      |||         |||   :::
Db     16270 CATGCGCCTCCACCGCGTCGATGTCCTCGGGCGACAGCCGGCACGACGAGAGCGCCTGTT 16211

Qy      1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
                   |||   |||      |||            |||:::|||
Db     16210 GGATCACCCGCTGCTGGGCAGGGCCGTTCGGCGCCGTCAAACCCTGGCTGCGACCGTCCT 16151

Qy      1981 ProArgProLeuVal---------ProProValSerGlyHis------------------ 1991
                   ||||||            ||||||      ||||||
Db     16150 GGTTCACCGCTGACCCACGGATCACCGCCAGCACACGGTCACCGTCGCGCCGCGCGTCCG 16091
```

```
Qy       1992 ------------AlaThrIleAlaArgThrPro------------AlaLysAsnLeuAla 2003
                          |||      :::|||||||||               :::   ||||||
Db      16090 ACAGCCGCTTCAGCACCAGCATCCCGCACCCCTCGGACCAGATGACGCCGTCAGCTCGCG 16031

Qy       2003 ------------------------------------------------------------ 2003

Db      16030 CAGAGAAGCTCTTGCAACGACCGTCCCGTGCAAGGGCCTTGAGGCGGCTGAACTCCACGA 15971

Qy       2004 ---------------------------ProHisHisAlaSerProAspProProAlaPro 2014
                                         ||||||   :::   ||||||         |||
Db      15970 ATCCCGCGGGTGTGGTCATCACCATCACCCCACCGGTCAAGGCCAGATCGCACTCGCCCT 15911

Qy       2015 ProAla---------------SerAlaSerAspProHisArgGluLysThrGlnSerLys 2029
                 |||                  |||:::   |||            |||
Db      15910 GACGCAGGGCCGTGCACGCCAGGTGCAGCGACACCAGCGACGACGAGCACGCCGTGTCCA 15851

Qy       2030 ProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSer 2049
              |||                           |||||||||            :::   :::
Db      15850 CCG------------------------TGATCGCTGGGCCCTGCAGCCCGAGCACGTACG 15815

Qy       2050 ProGluGlyValGluProValSerProValSer---SerProSerLeuThrHisAspLys 2068
              |||   |||:::         ||||||   |||      |||
Db      15814 CCACACGGCCTGAGAGCACGCTCGCCGCGCTCCCGGTGCCCCGGT--------------- 15770

Qy       2069 GlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgPro 2088
                    |||                     |||   |||                  |||
Db      15769 AGCCGTCCAACGCCGCCAGGTCGCTGCCGTAAAGAGCACCGTAGTCCGAGC------CCA 15716

Qy       2089 LysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgPro 2108
                    ||||||      :::            :::|||               :::|||
Db      15715 TCGAGCCCAGGTACACCCCGGTGGAGCTCTCGCTCAGCGCCGACGGGCGCACGCCAGCCC 15656

Qy       2109 LeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGly 2128
                 |||         |||:::   ||||||            ||||||         |||
Db      15655 GCTCCA---GTGCCTCCCACGCCGTCTCCA------------GCACCAGCCGCTGCTGGG 15611

Qy       2129 HisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThr 2148
                                |||                                    :::
Db      15610 GATCCATCGCCTGCGCCTCTCGGGGCG---------------------------ATATCC 15578

Qy       2149 ArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGly 2168
              |||   :::|||   :::   :::         ||||||   ::::::
Db      15577 CGAAGAAACCCGCATCGA---ACAGGTCGATGTCCCGCAGGAATCCACCCTCGCGCACGT 15521

Qy       2169 AlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProPro 2188
                 |||||||||                                       |||||||||
Db      15520 AGGTCTTGCCCA---------------------------------------CCGCCTCCG 15500

Qy       2189 AspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArg--------- 2205
              |||   |||      ::::::      |||   ::::::|||      |||
Db      15499 GATCGGGGTCGTAAATCGAAGGTGCATCCCAGCGCGCAGGGAAGGCCTCGATCGCATCTC 15440

Qy       2206 SerProGluProAsnLysThrSer------------------------------------ 2213
              ||||||   |||         |||
Db      15439 TCCCCTCCTCCAGGAGACGCCAGTAGTCTTCCGGCGTAGCGACGCCCCCCGGCAGCCGGC 15380

Qy       2214 ------------------------ValLeuGlyGlyGlyGluAspGlyIleGluProVal 2225
                                         |||   |||         |||
Db      15379 ACGCCATCGATACGATGGCGATGGGCCCTGAACGGAGATCCGAGTGGCTGGCCTGCAGCC 15320

Qy       2226 SerProProGluGlyMet------ThrGluProGlyHisSer------------------ 2237
              |||            |||      :::   |||      :::
Db      15319 TCTGCAACTCTTTCAATGCACGGGTCAACGCCTCACGAAGCGTAGGAGAGGCGGCGCTGC 15260

Qy       2238 ---------ArgSerAlaValTyrProLeuLeuTyrArg------------------Asp 2248
                       ||||||:::   :::   ::::::   :::
Db      15259 TATTCATGGAGATCCTCTGAGTTCGAAATCATTTGCAAGGCTTGCAAGGCCAGCTGCTCG 15200

Qy       2249 GlyGluGlnThrGluProSerArgMetGlySer--------------------------- 2259
                    ::::::      |||   ||||||:::
Db      15199 TCGGCGAGATCTTCGAAATCGGAAATGGGAACCGGAACCTCGACGGATGTTTCGGACGGC 15140

Qy       2260 ---LysSerProGlyAsnThrSerGlnProProAlaPheSerLysLeuThrGluSer 2278
                 :::|||      :::      :::   |||         ::::::|||||||||
Db      15139 CCCCGCAGCTCCAAGAGCTTGGGGAGGAGGCCTGAACGACGCACGAGACTGACGGAGAGG 15080

Qy       2279 AsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsn 2298
                 |||         :::            ::::::         :::   :::
Db      15079 CTGAGCATGAACGCGCGAATTTCCTCGTCGGACATGGAGTCCGCTGTACGCAGCAGATTC 15020

Qy       2299 ArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAla 2318
              |||      |||      |||      :::               |||   |||
Db      15019 CGGCGCTCACCGTTTCCGAACTTTGACAGGAGGAAC---------TTCGCGATGTGCGTG 14969

Qy       2319 IleThrGly------------------ThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
                    |||                     |||   :::         |||:::|||
Db      14968 GCAGATGGATGGTCGAAAAGCAAAGTGGCAGGCAAGTTGCTTCCGACCAGCAAGGCGAGC 14909

Qy       2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
              :::         ||||||   |||
Db      14908 CGATTGCGGACTTCGACGGCCATG------------------------------------ 14885

Qy       2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
                                ||||||   |||   |||         :::
Db      14884 ---------------AGCGAATCCATTCCGAGCTCCCTCAAGGGCTGGGCCGTAGGAATC 14840

Qy       2373 AlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHis 2392
              ||||||         |||         |||   |||||||||
Db      14839 GCGTCA---------CCGCGCTGAAGCCCGAGCACCGCGGCGACCTCGGCCCGCACAAGG 14789

Qy       2393 ThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArg 2412
              :::::::::                  :::         ||||||            |||
Db      14788 TCGATGAGCGAGCTC------------AGGCGCTCCGCCTCCGGCAGCGCCGAGAGGCGC 14741

Qy       2413 LysAlaLysSerProAlaProGlyLeuAlaSerGlyAsp------ArgProProSerVal 2430
                             |||      ||||||:::   |||      |||   |||
Db      14740 TCGCGGAGCGTCGACGCCTCCTTCCTCGCGGCAGAGGACGCCTTGCGCAGCCCGGGGCGC 14681

Qy       2431 SerSerValHisSerGluGly---------------------------------AspCys 2439
              :::|||      |||                                          |||
Db      14680 ACCAGCGCGCGAAGCAGCGCCGGCAGCTCGCCGCTGGACTCCAGCCCACGCTGCAGCTGC 14621

Qy       2440 AsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySer 2459
                 |||                  |||:::      :::|||   |||:::|||   |||
Db      14620 GCGAGATCCAGGTGCACTGGGACCAGGCTCGCTTCAGAGCGTGAGAGCGCGGCGTCCAGC 14561

Qy       2460 ThrProPheProTyrAsn-ProLeu------------------------------IleMe 2469
                       |||   ::: |||:::                                 ::
Db      14560 AGAGCGAGGCCCTCGTCGACCGATATCGGCACGAGCCCTGCGCGCCTGATGCGCGACAGT 14501

Qy       2469 tArgLeuGlnAlaGlyValMetAlaSerProProPro------ProGlyLeuProAla-- 2486
              :|||   :::   |||   ::::::||||||   |||      |||      |||
Db      14500 TCGGCCTCGCCCAGGTGCGCTGTCATCCCCACGCCCGCTTGCGCCCACAACCCCCACGCC 14441

Qy       2487 -GlySerGlyProLeuAla 2492
               ||||||   |||   |||
Db      14440 AGGCTCGTCGCCGCGAGCC 14422

RESULT 4
US-08-372-652-9
; Sequence 9, Application US/08372652
; Patent No. 5932699
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, David
;    APPLICANT:  Seol, Wongi
;    APPLICANT:  Choi, Hueng-Sik
;    TITLE OF INVENTION:  RETINOID X RECEPTOR-INTERACTING
;    TITLE OF INVENTION:  POLYPEPTIDES AND RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/372,652
;       FILING DATE:  13-JAN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00786/246001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617/542-5070
;       TELEFAX:  617/542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1922 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-372-652-9

Alignment Scores:
Pred. No.:              1.08e-24      Length:         1922
Score:                  718.00        Matches:        212
Percent Similarity:     48.81%        Conservative:   76
Best Local Similarity:  35.93%        Mismatches:     214
Query Match:            5.43%         Indels:         88
DB:                     2             Gaps:           25

US-09-522-753-5 (1-2517) x US-08-372-652-9 (1-1922)

Qy       1975 ProSerLysGlySerGluProArgProLeuVal--------------------------- 1985
              ||||||      ::::::|||         |||
Db        209 CCAAGTGGCAAGGCCCAGCCTCATGCCTCAGTAGTGTATTCTGAGGCTGGGAAAGATAAA 268

Qy       1986 ------ProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn------ 2001
                    ||||||   |||   :::            ||||||      |||
Db        269 GGGCCTCCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCCGAGGGAAGACTACCATT 328

Qy       2002 LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021
                 |||   :::                              ||||||      |||
Db        329 ACTGCAGCTAACTTCATAGACGTGACCATCACCCGGCAAATTGCCTCGGACAAGGATGCG 388

Qy       2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeu 2041
                          :::||||||         |||                     ||||||
Db        389 AGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT------------------AGTAGCTTG 430

Qy       2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061
                    |||            :::   :::   :::|||   :::||||||   ||||||
Db        431 TCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCCCGCCAGCTCA 490

Qy       2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081
              |||:::         :::|||                  :::   :::      :::
Db        491 CCTGCACCACCCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAAGGCAAATCAA 550

Qy       2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAla 2101
                 |||   |||         :::      ||||||
Db        551 GCAGAAAATGAGTCCACTCGACAGTATGAAGGTCCA------------------------ 586

Qy       2102 HisLeuProHisLeuArg---------------ProLeuProGluSerGlnProSerSer 2116
                 |||   |||   |||               |||   |||:::   ||||||
Db        587 ---CTGCATCATTATCGGTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTG 643

Qy       2117 SerProLeuLeuGlnThr------AlaProGlyValLysGlyHisGlnArgValValThr 2134
                 |||      |||:::               |||            |||::::::|||
Db        644 CCCCCATCTTCCCAGTCAGAGGGAATGGACAGGTGCCCAGGACCCATCGACTGATCACA 703

Qy       2135 LeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGln 2154
              ||||||   ||||||   ::::::||||||||||||:::   |||:::      |||
Db        704 CTTGCTGACCACATCTGTCAAATTATCACACAAGATTTGCTAGAAAT------CAAGTT 757

Qy       2155 LeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAla------SerCysPro 2172
                 |||   |||      :::         :::   |||   |||      |||   |||
Db        758 CCCTCGCAGCCTTCTACTTCTACATTCCAAACTTCACCATCTGCTTTGTCATCCACACCT 817

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAsp--------- 2189
              |||         |||         |||      |||   |||
Db        818 GTA---------AGAACTAAAACCTCAAGCCGCTACAGCCCAGAATCACAGTCTCAGACT 868

Qy       2190 ------HisGlyAlaProAla---ArgGlySerProHis---------------SerGlu 2201
                    |||      |||      |||   ||||||                  |||
Db        869 GTCTTGCATCCCAGACCAGGTCCTAGAGTCTCTCCAGAAAATCTTGTGGATAAATCCCGG 928

Qy       2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGly 2221
              |||   :::            |||   ::::::   :::               :::
Db        929 GGAAGCAGGCCTGGAAAATCTCCAGAGAGGAGTCATATC---------CCATCAGAGCCC 979

Qy       2222 IleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241
                 ||||||:::|||||||||:::|||               |||   :::
Db        980 TATGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTG------CATGAGAAGCAGGACAGC 1033

Qy       2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261
                    ||||||   :::   |||                  |||      |||:::|||
Db       1034 ATGTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAGTGATTCTCGATCA 1093

Qy       2262 ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAla 2281
              ||||||:::   |||      |||:::||||||:::||||||   ||||||   |||
Db       1094 CCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAGCACATCACCC 1150

Qy       2282 MetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGlu 2301
              |||||||||||||||||||||||||||   :::|||||||||:::            :::
Db       1151 ATGGTTAAATCAAAGAAACAGGAAATTTTCGTAAGTTGAACTCTTCTGGTGGAGGTGAC 1210

Qy       2302 ProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGly 2321
                 :::         :::||||||||||||||||||||||||:::||||||:::|||
Db       1211 TCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGCAGTTACCACA 1270

Qy       2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341
              :::|||   ::::::   ||||||   :::      :::   ||||||   |||:::|||
Db       1271 TCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGT---AACCTTGGT 1327

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu------ 2359
              ||||||   ||||||||||||||||||||||||   :::|||         |||
Db       1328 CTAGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGAAGATCAT 1387

Qy       2360 ------------SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAla 2375
                          |||   |||:::                  |||   :::||||||
Db       1388 GGTGTTGTCATGTCCCATCCTGTG---------GGCATTATGCCTGGTAGTGCCAGCACC 1438

Qy       2376 SerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThr 2395
              |||:::               :::|||:::::::::   |||   |||
Db       1439 TCAGTG---------------GTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCA 1483

Qy       2396 SerProGlyGlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAla 2414
              ||||||      |||      |||   |||:::      :::   :::|||||||||:::
Db       1484 TCACCTCATGCAGGAGTATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCT 1543

Qy       2415 LysSerProAlaProGlyLeuAla------SerGlyAspArgProProSerValSerSer 2432
              |||||||||   ||||||   :::            :::||||||   ||||||||||||
Db       1544 AAATCTCCTATTCCTGGGCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCT 1603

Qy       2433 ValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAsp 2452
              ||||||||||||||||||   :::|||:::||||||               |||||||||
Db       1604 GTGCATTCAGAAGGTGATTACCACAGGCAGACACCA------GGATGGGCATGGGAAGAT 1657

Qy       2453 ArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGln 2472
              ||||||||||||   |||||||||   ||||||||||||||||||   :::|||:::
Db       1658 CGGCCCTCTTCAACAGGTTCTACTCAGTTCCCTTACAACCCTCTGACCATACGGATG--- 1714

Qy       2473 AlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu--- 2491
                       ::::::|||   ||||||      :::                  :::
```

```
Db       1715 ---------CTCAGCAGTACACCACCTACACAGATCGCATGCGCCCCATCTGCCATCACC 1765

Qy       2492 ---AlaGlyProHis---------HisAlaTrpAspGluGluProLysProLeuLeuCys 2507
                 |||   ||||||               |||:::   ||||||   |||||||||
Db       1766 CAAGCAGCTCCACATCAACAGAACCGCATCTGGGAGAGGGAGCCTGCCCCGCTCCTCTCA 1825

Qy       2508 SerGlnTyrGluThrLeuSerAspSerGlu 2517
              :::||||||||||||||||||||||||:::
Db       1826 GCGCAGTATGAGACACTGTCTGATAGTGAC 1855

RESULT 5
PCT-US95-16311-9
; Sequence 9, Application PC/TUS9516311
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, David
;    APPLICANT:  Seol, Wongi
;    APPLICANT:  Choi, Hueng-Sik
;    TITLE OF INVENTION:  RETINOID X RECEPTOR-INTERACTING
;    TITLE OF INVENTION:  POLYPEPTIDES AND RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/16311
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,652
;      FILING DATE:  13-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/246001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1922 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
PCT-US95-16311-9

Alignment Scores:
Pred. No.:               1.08e-24      Length:         1922
Score:                   718.00        Matches:        212
Percent Similarity:      48.81%        Conservative:   76
Best Local Similarity:   35.93%        Mismatches:     214
Query Match:             5.43%         Indels:         88
DB:                      5             Gaps:           25

US-09-522-753-5 (1-2517) x PCT-US95-16311-9 (1-1922)

Qy       1975 ProSerLysGlySerGluProArgProLeuVal--------------------------- 1985
              ||||||      ::::::|||         |||
Db        209 CCAAGTGGCAAGGCCCAGCCTCATGCCTCAGTAGTGTATTCTGAGGCTGGGAAAGATAAA 268

Qy       1986 ------ProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn------ 2001
                    ||||||   |||   :::            ||||||      |||
Db        269 GGGCCTCCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCCGAGGGAAGACTACCATT 328

Qy       2002 LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021
                 |||   :::                              ||||||      |||
Db        329 ACTGCAGCTAACTTCATAGACGTGACCATCACCCGGCAAATTGCCTCGGACAAGGATGCG 388

Qy       2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeu 2041
                          :::||||||         |||                     ||||||
Db        389 AGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT------------------AGTAGCTTG 430

Qy       2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061
                    |||            :::   :::   :::|||   :::||||||   ||||||
Db        431 TCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCCCGCCAGCTCA 490

Qy       2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081
              |||:::         :::|||                  :::   :::      :::
Db        491 CCTGCACCACCCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAAGGCAAATCAA 550

Qy       2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAla 2101
                 |||   |||         :::      ||||||
Db        551 GCAGAAAATGAGTCCACTCGACAGTATGAAGGTCCA------------------------ 586

Qy       2102 HisLeuProHisLeuArg---------------ProLeuProGluSerGlnProSerSer 2116
                 |||   |||   |||               |||   |||:::   ||||||
Db        587 ---CTGCATCATTATCGGTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTG 643

Qy       2117 SerProLeuLeuGlnThr------AlaProGlyValLysGlyHisGlnArgValValThr 2134
                 |||      |||:::               |||            |||::::::|||
Db        644 CCCCCATCTTCCCAGTCAGAGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACA 703

Qy       2135 LeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGln 2154
              ||||||   ||||||   ::::::||||||||||||:::   |||:::      |||
Db        704 CTTGCTGACCACATCTGTCAAATTATCACACAAGATTTTGCTAGAAAT------CAAGTT 757

Qy       2155 LeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAla------SerCysPro 2172
                 |||   |||      :::         :::   |||   |||      |||   |||
Db        758 CCCTCGCAGCCTTCTACTTCTACATTCCAAACTTCACCATCTGCTTTGTCATCCACACCT 817

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAsp--------- 2189
              |||         |||         |||      |||   |||
Db        818 GTA---------AGAACTAAAACCTCAAGCCGCTACAGCCCAGAATCACAGTCTCAGACT 868

Qy       2190 ------HisGlyAlaProAla---ArgGlySerProHis---------------SerGlu 2201
                    |||      |||      |||   ||||||               |||
Db        869 GTCTTGCATCCCAGACCAGGTCCTAGAGTCTCTCCAGAAAATCTTGTGGATAAATCCCGG 928

Qy       2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGly 2221
              |||   :::            |||   ::::::   :::               :::
Db        929 GGAAGCAGGCCTGGAAAATCTCCAGAGAGGAGTCATATC---------CCATCAGAGCCC 979

Qy       2222 IleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241
                 ||||||:::|||||||||:::|||               |||   :::
Db        980 TATGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTG------CATGAGAAGCAGGACAGC 1033

Qy       2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261
                    ||||||   :::   |||                  |||      |||:::|||
Db       1034 ATGTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAGTGATTCTCGATCA 1093

Qy       2262 ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAla 2281
              ||||||:::   |||      |||:::||||||:::||||||   ||||||   |||
Db       1094 CCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAGCACATCACCC 1150

Qy       2282 MetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGlu 2301
              |||||||||||||||||||||||||||   :::|||||||||:::            :::
Db       1151 ATGGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGGTGGAGGTGAC 1210

Qy       2302 ProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGly 2321
                 :::         :::||||||||||||||||||||||||:::||||||:::|||
Db       1211 TCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGCAGTTACCACA 1270

Qy       2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341
              :::|||   ::::::   ||||||   :::      :::   ||||||   |||:::|||
Db       1271 TCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGT---AACCTTGGT 1327

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu------ 2359
```

```
                ||||||   ||||||||||||||||||||||||   :::|||    |||
Db         1328 CTAGAAGACATCATCAGAAAGGCTCTCATGGAAGTTTTGATGATAAAGTTGAAGATCAT 1387

Qy         2360 ------------SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAla 2375
                            |||   |||:::                  |||   :::||||||
Db         1388 GGTGTTGTCATGTCCCATCCTGTG---------GGCATTATGCCTGGTAGTGCCAGCACC 1438

Qy         2376 SerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThr 2395
                |||:::               :::|||:::::::::   |||   |||
Db         1439 TCAGTG---------------GTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCA 1483

Qy         2396 SerProGlyGlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAla 2414
                ||||||      |||      |||   |||:::      :::   :::|||||||||:::
Db         1484 TCACCTCATGCAGGAGTATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCT 1543

Qy         2415 LysSerProAlaProGlyLeuAla------SerGlyAspArgProProSerValSerSer 2432
                |||||||||   ||||||   :::            :::||||||   ||||||||||||
Db         1544 AAATCTCCTATTCCTGGGCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCT 1603

Qy         2433 ValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAsp 2452
                ||||||||||||||||||   :::|||:::||||||               |||||||||
Db         1604 GTGCATTCAGAAGGTGATTACCACAGGCAGACACCA------GGATGGGCATGGAAGAT 1657

Qy         2453 ArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGln 2472
                ||||||||||||   |||||||||   ||||||||||||||||||   :::|||:::
Db         1658 CGGCCCTCTTCAACAGGTTCTACTCAGTTCCCTTACAACCCTCTGACCATACGGATG--- 1714

Qy         2473 AlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu--- 2491
                         ::::::|||   ||||||      :::                  :::
Db         1715 ---------CTCAGCAGTACACCACCTACACAGATCGCATGCGCCCCATCTGCCATCACC 1765

Qy         2492 ---AlaGlyProHis---------HisAlaTrpAspGluGluProLysProLeuLeuCys 2507
                   |||   ||||||               |||:::   ||||||   |||||||||
Db         1766 CAAGCAGCTCCACATCAACAGAACCGCATCTGGGAGAGGGAGCCTGCCCCGCTCCTCTCA 1825

Qy         2508 SerGlnTyrGluThrLeuSerAspSerGlu 2517
                :::||||||||||||||||||||||||:::
Db         1826 GCGCAGTATGAGACACTGTCTGATAGTGAC 1855

RESULT 6
US-09-902-540-1090/c
; Sequence 1090, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1090
;   LENGTH: 14462
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-1090

Alignment Scores:
Pred. No.:              7.98e-23       Length:        14462
Score:                  698.00         Matches:       652
Percent Similarity:     32.50%         Conservative:  301
Best Local Similarity:  22.24%         Mismatches:    1028
Query Match:            5.28%          Indels:        964
DB:                     4              Gaps:          143

US-09-522-753-5 (1-2517) x US-09-902-540-1090 (1-14462)

Qy           44 HisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGlnProGlnArg 63
                |||   ||||||                  |||   |||:::         |||
```

```
Db        14437 CACCCCAGCCGCTCCAGCGGGTAGCTCGCGTCCAGCGGCACGTACACGCCCCCCGCCTTC 14378

Qy           64 ArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSer-GlnGluLeuHi 83
                |||   ||||||   |||:::               |||||||||||| :::   :::
Db        14377 AGAATCCCCAGCGCGCTCACCACCCAC------TCCAACGAGCGCTCCGCGCACAATCCC 14324

Qy           83 sLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIl 103
                 |||   |||
Db        14323 ACTCGCACCTCGGGGC-------------------------------------------- 14308

Qy          103 eGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgProSerProLe 123
                                                                 |||   |||
Db        14307 --------------------------------------------------CCACTCCCATT 14297

Qy          123 uLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThrGl 143
                    :::   |||   |||   |||            :::      |||:::
Db        14296 CCCCTCAGGTGGTGCGCCAACTGGTTTGCTCGGCGATTCAACGCCTCGTAACTCACCTGC 14237

Qy          143 yLysLeuGluProValSerProProSerProProHisThrAspProGluLeuGluLeuVa 163
                    |||            ||||||   |||      :::   |||   |||:::
Db        14236 TGCCCTGCCCACTCCACTGCCACCGCGCCCGGGGTCCTCTCCACCTGCGCTTCAAACAGC 14177

Qy          163 lProPro-ArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGlu- 182
                 |||||| |||                           :::   ||||||   |||
Db        14176 TCCGCCAGCGA-------------------------CGCCTCCCTCGGGTACTCCGCTCGC 14141

Qy          183 -----IleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuG 201
                           ::::::   |||            |||   :::   |||:::||||||
Db        14140 CTCCCGCTCCACTCCTCCACCAGGACGCGCCGCTCCTCCACCGACATCAACGACAGCTCG 14081

Qy          201 luGluGluAlaAlaLysProProGluProGluLysProValSerPro------------- 216
                  :::   |||:::            |||      |||   :::|||
Db        14080 CCCAGGC-GCGTCTCCGGATTCGCGACCGCGGCCTCCAGCAGCACCGACATCTGCCTCAG 14022

Qy          217 --------ProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluA 234
                        ||||||      :::   :::|||||||||:::   :::
Db        14021 CAGCCGCTCCGCCGTCTTC-GCCTCGAACAGGTCGCTGTTGTAATTG------------- 13976

Qy          234 snArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluL 254
                           |||   |||      |||   |||:::            ||||||||||
Db        13975 -----ACTGACGCGGCGGCACC-TTCCGGCAGCTCGACGACGAACATCGACAGGTCGAAC 13922

Qy          254 euProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnG 274
                ||         |||                                    :::::::::
Db        13921 TTCGACGTCTGAATCCG-----------------------------------CGCGTCCATC 13895

Qy          274 lnAlaMetArgLysLysLeuIle---LeuTyrPheLysArgArgAsnHisAlaArgLysG 293
                        |||:::   |||      |||:::|||      |||         :::   |
Db        13894 ACCGTCAGCGACAATCCCTGTCGCAGCTTCACTTCGGGCACAGGCGTGTTCTGAAGTGTC 13835

Qy          293 ln------TrpLysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluL 311
                ||         ||||||:::         :::            ::::::   |||   :
Db        13834 AGGGAGACTTGAAACAACGGGCTGCGACTCAGGTCCCGCTCCGGCTGCAATGCCTCCACC 13775

Qy          311 ysLysValGluArgIleGlu-----------------------------AsnAsnProA 321
                ::   :::||||||                                       :::   |
Db        13774 AGCTTCTCGAACGGGACGTCCTGGTGCGCGTAGGCGCCCAGGGTCACCTCACGCACCTGC 13715

Qy          321 rgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleA 341
                ||:::         |||   |||:::      |||         |||         |||
Db        13714 GCCAACATGTCCCGGAAGGAAAGCTCGGGACGTACGCGAGCCCTCAGCACCAGGGTATTG 13655

Qy          341 rgLysGlnArgGluLeuGlnGluArgMet---------GlnSerArgValGlyGlnArgG 358
                  :::      :::         |||:::            :::|||   |||:::
Db        13654 GCGAAGTTGCCGATGAGCCCCTCCGTCTCGGCCTGCGTGCGACCCGCGATGGGGAAGCCC 13595

Qy          358 lySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspG 378
                  :::   :::   :::               :::|||   :::
Db        13594 ACGCTGATGTCGTCCTGCCCCGAGAGCCGCCACAACACCGTCTG---------------- 13551

Qy          378 lyLeuSerGluGlnGluAsnLeuGluLysGln------------MetArgGlnLeuAlaV 394
                  :::      |||::::::   |||::::::               |||
Db        13550 --ATACGCCGCCAGCAGCACCATGAACGGCGTGGCCCCTTCTCGCCGCGCGGTCGCCTCC 13493
```

```
Qy        394 alIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnG 414
                :::||||||                     ||||||::::::
Db      13492 ACTTGCCGCCA---------------------CAGCGCCTTCGG---------------- 13470

Qy        414 lyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerG 434

Db      13470 ------------------------------------------------------------ 13470

Qy        434 luGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuI 454
                                                    |||
Db      13469 -----------------------------------CCACAT------------------- 13464

Qy        454 leAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrL 474

Db      13463 --------------------------------------------------------CTGC 13460

Qy        474 ysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerG 494
                   :::         |||   |||   |||   :::         ||||||:::   |
Db      13459 CGCTGGATTCCACCTCGAAAGGTCTGCGCCGCGGGGCGAGGCCTGTCCGTGGGAGGTCC 13400

Qy        494 lnGlnGlnGlnGlnGlnGlnGln------------------------------------- 502
              ||:::   ::::::   :::|||
Db      13399 AGGAATCGCGGCGCGCCCGCCAGCACCTCGCGCCAGTAGCGAACCTGCGATTCCAACACC 13340

Qy        503 -----------------------------------------------------GlnGlnG 505
                                                                   |||   :
Db      13339 TCGTCCTTCAACCACTCGCGCTGCCAGAGCGCGTAGTCCGCGTACTGCACGGGCAGCTCC 13280

Qy        505 lnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluL 525
              ::|||      :::||||||:::   |||            |||:::   :::      |
Db      13279 GGCAATGGTGACGGCAGCCCTTG---CGAGAACGCGTCGTAGAGCGCCGCCAACTCCCGA 13223

Qy        525 ysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsnAspLysGluAspL 545
              ||:::                                 ::::::         ||||||:
Db      13222 AGCAGCACGCTCGCGGACCAACCGTCCGAGACAATGTGATGCAACAGCAGCACGAGGACG 13163

Qy        545 euLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValA 565
              ::         :::|||      |||:::|||
Db      13162 TGCG-CCTGCTCGGACTGCTTCAACAGCAGGG---------------------------- 13132

Qy        565 laSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyArgIleThrArgS 585
                      |||         |||:::::::::   |||   :::||||||:::   |||:
Db      13131 --------GGCGCAGCAGCGGCCCACGCGCGAGCTCGAATGCGCGGTCGACTTCCGCGCG 13080

Qy        585 erMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGlnSerAlaGluLeuA 605
              ::            |||   |||:::   :::      |||      :::   |||
Db      13079 CA---CCCGGTGCCGCACC-TCGGACTCATCCTGGGCGCCCAGCTCCACCACGGAGAGGT 13024

Qy        605 laSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysL 625
                :::   :::      |||:::   |||||||||               :::   |||
Db      13023 GGACCGGCGAT---GGCGAAGCGATGCGCTGGACAGGTGCTCCGCCTTCATCCGGGAAG- 12968

Qy        625 ysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysT 645
                                              |||:::   |||            :::|
Db      12967 --------------------------------GCCGTACGCAGGACTTCATGACGGCGCA 12940

Qy        645 hrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAsp---- 663
              ||   |||:::                        |||   |||         :::
Db      12939 CCAGCTCCGAG------------------------AAGGCTCGCTCCAAGGCAGAAACAT 12904

Qy        664 --------GluIleLeuGlnGln-HisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                               :::    |||   ::::::      :::   |||   |||
Db      12903 CCAGGGGGCCCTCCACCCGCAGCACATACGGTGCGTTGTACTCGGTACCGCCCGGCTTCA 12844

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAla---PheProProVal 699
                                   |||||||||         ||||||      ||||||
Db      12843 GCCGATCCAGGAACCACAGTCGCTGCTGCGCGAATGACAGCGGCACCCGGTCCCCCCGAG 12784

Qy        700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
                    |||            ||||||
Db      12783 GCATCGGACGCAGCGCGAGCAGCGTCTGACTTC--------------------------- 12751

Qy        720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluVal------------ProArgGlyGlu 735
              |||:::   |||      |||                              ||||||
Db      12750 GAGTCGCCTGCTCCACCTGCTGGGCAAAGGACCGCAAGGTGGGCGCCTCGAAGAGGGCGC 12691

Qy        736 CysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHis 755
                 :::||||||                                    |||   ||||||
Db      12690 GTAGCGGGACCTCTACGC------------------------------CCAGGACCGCAC 12661

Qy        756 ThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAla 775
                    |||                        |||   |||         :::
Db      12660 GCACACGCG------------------------CCACCACCTGCGTGGCCAGTAGCGAAT 12625

Qy        776 AspGlyPro---------------ProProGlyProProThrProProArgArgThrSer 790
              |||   |||               |||||||||||||||   ||||||   |||
Db      12624 GACCACCCGACGCGAAGAAGTCGTCCTCCCGGCCCACCCGCTCCACCCGCAAGA------ 12571

Qy        791 ArgAlaProIleGlu---------------------------ProThrProAlaSerGlu 801
              |||   |||   :::                           |||   |||
Db      12570 CGTGGGCCCAGACAGCGGCCAACGACTCTTCCATGGGGGAAGCCGGCGCCTCATAGGCGG 12511

Qy        802 AlaThrGly---AlaProThr---------------------------ProPro------ 809
                    |||   |||||||||                           ||||||
Db      12510 GAGCCGGGCCACGCGCCGACACGTTCGGCACGGGCAGCGCGCGCCAGTCCACCTTGCCAT 12451

Qy        810 ---------------------ProAlaProProSerProSerAlaPro------------ 818
                                   ||||||||||||:::|||:::||||||
Db      12450 TGGGCGTCAGCGGCAGGGCGTCCAGCACCACCAGCGCCGACGGCACCATGTGCCCAGGAA 12391

Qy        818 ------------------------------------------------------------ 818

Db      12390 GCCGCTGCTGGAGGAAGCCCCTGAGGGCCGCGATGTCCGGAGCCGCCCTGGCCGTACGAA 12331

Qy        819 ---------ProProVal---------------------------------ValProLys 824
                       ||||||                                       |||:::
Db      12330 CGGCGTAGCCCACCAGCCGATTGTCTCCAGGCACCTCTTCTCGCGCGACCACCACCGCGT 12271

Qy        825 GluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLys 844
                    :::   :::         |||   ||||||:::      |||      :::
Db      12270 CCTGGACGTCGGGATGCGCTCTGAGCACCGCCGCCACTTCT-CCGGGCTCCACCCGGAAC 12212

Qy        845 ProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSer 864
              |||                           |||   |||:::      |||:::   |||
Db      12211 CCT------------------CTCACCTTCACCTGGAAGTCCGTCCGCCCGATGAACTCC 12170

Qy        865 Glu------CysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAla 882
                                               |||   :::         |||   |||
Db      12169 AGTCGCCCATCCCCCAACCACCGCACCTTGTCCCCCGTGCGGTACATCCTCGCTCCAGCC 12110

Qy        883 AlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArg 902
              :::               |||      ::::::                  |||||||||
Db      12109 TCCATGCTGAACGGGTCCGGCACGTACCGCTCCGCCGTC---------AACTCCGGGCGC 12059

Qy        903 AlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSer 922
                 |||      :::   :::   ||||||      ::::::|||      |||   :::
Db      12058 CATACGTACCCTCGCGTGACCTGCGCCCCTCCCACGAACAACTCTCCCCCCACGCCCACC 11999

Qy        923 AlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer-ProArgPr 942
                                         |||         ||||||   |||    |||
Db      11998 ---------------------------GGCACAGGACGCAGGCTCCGGTCCAGCACGTAA 11966

Qy        942 oSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------Pr 960
                  :::      |||:::      |||         |||:::|||   :::      ||
Db      11965 ACCCGTGTGTTGCCCATCGGCCGTCCCA---------GCGGCACCGTCGCGGGTGACTCC 11915

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal----- 978
              |||||||:::                  |||         ||||||
Db      11914 GCTCGACGTG---------------GCAGCTCCACTCGCCCCGCCAGCACACCCACCGTC 11870

Qy        979 ----------------------------------ThrLysValHisGluProProArgGl 987
                                                |||:::         ||||||
Db      11869 GTCTCCGTCGGGCCGTAGTGGTTGTGCACCTCGCACCCGGGCGCCAGCGCCGCCACCTGC 11810

Qy        987 uAspAlaAlaProThrLysProAlaProProAlaProProPro----------------- 1001
```

```
                |||      |||:::|||   |||      |||||||||
Db    11809 TCCAGCAGGCTCCACGAGGCCG---CCTCTCCTCCCACCACCAGCCGCTTCGTCGGCAAC 11753

Qy     1002 -------------------------------ProGlnAsnLeuGln------ProGluSe 1009
                                           |||:::      :::      |||   ::
Db    11752 ACCTGCCCGGGCGCGGCGCCACTCATCAGCGCCGCGAGGTGCGAGGGCACCACCTTCAAA 11693

Qy     1009 rAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAlaProPr 1029
            :      |||         ||||||:::|||            |||            ||
Db    11692 CACTCCACCCCATGCCGCTGGAAGTACTCCTCCACTCCCGCCGCGCTGCTTGCCCTCTCC 11633

Qy     1029 oAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTr 1049
            ||||         |||   |||                  |||         ||||||::
Db    11632 TGCGTCAGCACATGCA---GCAGTCCCCCCGTCCACAGCGCCGCGTACAGCACCGTGTTT 11576

Qy     1049 pThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAl 1069
            :            |||   |||   ||||||            ::::::|||
Db    11575 C------------CCAGGTCCGCCGCCACCGTCGACACCAGCGCGAAGCTCGTGCACTCC 11528

Qy     1069 aProAspProSerAlaPheSerTyrAlaProProGly----------------------- 1081
             |||   |||:::|||   |||      |||
Db    11527 GCCAACCCCAGCCGCGCATTCGCCGCTTCCACGTAGTGGGCAAGCTGTCCGTGTGTGACC 11468

Qy     1081 ------------------------------------------------------------ 1081

Db    11467 GCCACGCCCTTGGGCCGCCCCGTGCTTCCGGATGTGAACAGCACGTACGCCAGTTGCTCG 11408

Qy     1082 ----HisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg----- 1098
                   |||||||||      |||         |||            ||||||
Db    11407 GGTTGCACCTCTACCCGCGGTGCTTCCGTCGGCTGCGCCTCGAGCCGCGCCGCGTCTCCA 11348

Qy     1099 -------ProPro---------------------------------------ThrIleSe 1103
                   ||||||                                             ||
Db    11347 TCCATCTCCACCCGGCGCGCGGACAGTCCCTCGAAAGTCGCCGCATGCCGCGCCTGCGTC 11288

Qy     1103 rAsnProPro------ProLeuIleSerSerAlaLysHisProSerValLeuGluArgGl 1121
            |   ||||||      ||||||:::      |||:::         :::|||
Db    11287 ACCACCACCGGCGCGGCCACTTCTTCCACCAGCGCGCGCAGGCGGGATTCTG-------- 11236

Qy     1121 nIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAl 1141
                      :::::::::|||   :::   :::      :::|||   ||||||
Db    11235 ----------GTTGCGAGGGGTCCAGCGGCACGTACGCGCCTCCCGCCTTCAGAATCCCC 11186

Qy     1141 aLysAlaPro---------------ValGlyProValThrMetGlyLeuProLeuProMe 1156
                :::|||                     |||   |||      |||   ||||||
Db    11185 AGCATCGCCACCACCCACTCCACGCCCCGCTCCAGGCACAGCGCCACTGGCACTTCCGCG 11126

Qy     1156 tAspProLysLysLeuAlaProPheSerGlyValLysGlnGlu---------GlnLeuSe 1173
                |||            ||||||
Db    11125 CCCACCC------------CCATTCCCCTCAACTGGTGCGCCACCTGGTTGGCCTGCGCC 11078

Qy     1173 rProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSe 1193
             |||   |||            |||               |||   |||      |||||
Db    11077 TCCAACTGGCGGTACGTCAGCTCCCGCTCCTCGCACCGCACCGCGGGCGCCTCTGGCGTC 11018

Qy     1193 rValLeuArgGly-------ThrAlaLeuGlySerValProGlyGlySerIleThrLysG 1211
            |   |||||||||       |||   :::   |||:::|||      |||:::   :::
Db    11017 CGCTCTGCGTGGTGCTCGACACACCGGTGGACAGCCTCCCACCGGAGAGCCTGGGCCGAG 10958

Qy     1211 lyIlePro------------------SerThrArgValProSerAspSerAlaIleThrT 1225
              :::|||                  :::   |||:::                     |
Db    10957 CCGTTCCATTCCACCAGCACCTGATGACGCTCCGCATCGGTCAGCAG---------CGGT 10907

Qy     1225 yrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrA 1245
            ||   |||            |||         |||            |||
Db    10906 ATCTGGGAGAGGCGCGATTCGGGCCGCGTGGTGAC------------GGCCTTCAGCACC 10859

Qy     1245 rgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAsp---------- 1261
              :::      |||      |||   |||      |||||||||   :::
Db    10858 TGCTCCAGGTAGAACACCATCCGCTCCGC------CGTGGCCGCCGTGAAGAGGTCGGTG 10805

Qy     1262 --------------------------------SerLeuProLysGlyHisValIleTyrG 1271
                                            |||||||||:::               |
Db    10804 CTGAACTGGAGCAAGCCCTCGAAACCCCGTGGAGTCTCCCCCAGGCTCAACGAAAGCTCG 10745

Qy     1271 luGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerL 1291
            ||      :::                           :::   |||   |||      :
Db    10744 AAGCGAGCCGC---------------------------CTCCTTGTCAATCAACCCCATC 10712

Qy     1291 ysGluAspGly---ArgSerSerSerGlyProProHis---------------------- 1302
            ::      |||   |||               ||||||
Db    10711 GGCTTCACGGTCAGCGAGGGCAGCGACAACGGCCCCACCGGGGCGTTCTGCAACACGAAG 10652

Qy     1303 -----------GluThrAlaAlaPro------------------LysArgThrTyrAspM 1313
                       |||   |||||||||                  :::      :::::::
Db    10651 AGGACCTGGAAGAGCGCGCTGCGCCCCAGGTTCCGCTCGGGTTGCAGATGCTCCACGAGC 10592

Qy     1313 etMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyA 1333
            :::::||||||:::|||                                 ||||||
Db    10591 TTCTCGAAGGGAAGGTC---------------------------------CTGATG---- 10569

Qy     1333 rgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHis---------H 1350
                 :::   |||:::|||   :::      ||||||::::::|||
Db    10568 --CTCGTACGCCCCCAGCGT---GCGCTCACGCACCTGCGCCAGCAGTTCGCGGAAGGTG 10514

Qy     1350 isIleArg---------GlySerIleThrGlnGlyIlePro------------------- 1360
              :::|||         |||:::      ||||||:::
Db    10513 GCATGAGGCTGGACCTGGGCGCGGAGAACCAGGGTGTTGACGAAGAAGCCGATGAGCCCC 10454

Qy     1361 --ArgSerTyrValGluAlaGlnGluAsp---------------------TyrLeuArgA 1373
              |||   :::|||   |||      |||                        |||   |
Db    10453 TCCGTCTCTTCGTACCGGCGACCCGCGATGGGCGAGCCCACCAGCACGTCCTCCTGCCCC 10394

Qy     1373 rgGluAlaLysLeuLeuLysArgGluGlyThrPro------------------------P 1385
            ||   |||:::      :::   ||||||   |||                        |
Db    10393 GAGTAGCGCGACAGCAGGAGCTGGAAGGCCG-CCAGCAGCACCATGAAGGGCGTGGCTCC 10335

Qy     1385 roPro------ProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuG 1403
            |||||      |||   |||||||||      :::   |||      |||
Db    10334 CTCCGCATGGGCCAGTGCCGTCACGGCGTCGCTCATCGGCCGTGACAACCGCACCG---- 10279

Qy     1403 lyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGlyA 1423
                             |||   |||   |||      ||||||         |||
Db    10278 --------------GCACCCATGCACCGCGGT------GCGACAACGTCGCCGGCCTGGG 10239

Qy     1423 rgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuProLeu---- 1441
                             |||   ||||||   |||   |||
Db    10238 CTTGTCGGTGGGTAGCTCCAGGTGAGGAGGCGCGCCAGACAGCTGCTGCCGCCACCAGTC 10179

Qy     1442 --AlaProArgProLeuLysGluGlySer---IleThrGlnGlyThrProLeuLysTyrA 1460
              |||      |||   :::      |||   |||         ||||||
Db    10178 GAGCTGCGTCTCCAGCACGGCGCCGCTCAGCCATTGCCGCTGCCACACCGCGTAGTCCGC 10119

Qy     1460 spThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerP 1480
              |||      :::         :::::::::      :::|||         |||   |
Db    10118 GTACTGCACGGGCAGCTCCGGCAACCCGGGAGACCGTCCTTCGA---------GGAAGGC 10068

Qy     1480 roGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuG 1500
            ||   |||      ||||||      |||                  |||   :::
Db    10067 CGCATAGAGCGCGCCCACCTCGCGGACCA---------GCACGCCCAGCGACCAGCCATC 10017

Qy     1500 luArgAlaCysTyrGluGluSerLeu------LysSerArgProGlyThrAlaSerSerS 1518
              |||   |||                     ::::::||||||   ||||||:::::::
Db    10016 CGAGACGATGTGATGCAGCACCAGCACCAGGGCGTGCACGTCCGGCGACAGCCGCGACAG 9957

Qy     1518 er------------GlyGly------------------------------------SerI 1522
            ::            ||||||                                    |||
Db     9956 CCCCACGCGCACCAGGGGGCCGCGCTCCAGATGAAACGGCCGCAGGGCTGTTCGAGTCGC 9897

Qy     1522 leAlaArgGlyAlaProValIleValPro------------------------------G 1532
              |||||||||               |||
Db     9896 ACGCGCGCGGAATTCGGCCTCGCGCTCCACTTCCGGCAGCGTGCTCAAGTCCTCGAAGGC 9837

Qy     1532 luLeuGlyLysPro---------------------------------------ArgGlnS 1539
                 |||:::|||                                       ||||||:
Db     9836 GAGCTGGACGTCCATGGCCGGATGGATGACCTGGACAGGCTGGCCCTCGTGGGAGGCAAA 9777
```

```
Qy   1539 erProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerP 1559
          ::                  ||||||||||||   :::            |||   :::
Db   9776 CGTCGTGCGCA---ACGACTCATGGCGCGCCACC-AGCGCGGAGAACGCGCGCTGCAACG 9721

Qy   1559 roValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysA 1579
            :::|||      ||||||:::   |||      :::|||   :::   :::      |
Db   9720 CGCTCACATCGAGCGAGCCCTCGATCCGGAG-GACGATGGGGACGTTGTAGCTGGCACTG 9662

Qy   1579 laSerGlnAspArgLysLeuThrSerThrProArgGluIle---------------AlaL 1594
          ||      :::      |||         |||      :::                  :
Db   9661 CCCGGATGGAG------CTGGTCGATGAACCAGAGCCGCTGCTGGGCGAACGACACGGGC 9608

Qy   1594 ysSerProHisSerThrValProGluHisHisProHis---------------------- 1606
          ::      |||:::      |||            |||
Db   9607 AATGCGCCCATGCGAGGTACCCGGGTCAGCGCTGGCACGCTCGGCGCCGCCGCTTCCCGC 9548

Qy   1607 -----------------ProIleSerProTyrGluHis---------------------- 1613
                                    |||:::||||||
Db   9547 GCCGACTCCTCGATTCGCAGCGCGAGCCCTTCGAGCACAGGTGACGCGAAGAGCGTGCGC 9488

Qy   1614 -----LeuLeuArgGlyValSerGlyValAspLeuTyr---------------ArgSerH 1627
                  |||         ||||||   |||:::                  :::
Db   9487 AGGGGCAACTCCACGCCGAAAGCGGAGCGGATGTGCGCGATGAGCTGCGTGGCCAGCAAC 9428

Qy   1627 isIleProLeuAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaA 1647
            :::|||         :::      |||:::   |||      ||||||||||||
Db   9427 GAGTGCCCGCCCAACTCGAAGAAGTCAGCGTC---CGCGCTCACCCGCTCGATGCC---- 9375

Qy   1647 laAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProP 1667
                           :::|||      |||
Db   9374 -----------------CAGCACTTCACACCA---------------------------- 9360

Qy   1667 roTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleA 1687
                                                ::::::   |||   ::::::
Db   9359 --------------------------------------GATGATGCTCAGCTTCTGCTCG 9338

Qy   1687 snAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgA 1707
            |||                                                   |||
Db   9337 GTGAC---------------------------------------------CGTACGCGGC 9323

Qy   1707 laAspMetLeuArgGlyLeuSer---ProArg---------GluSerSerLeuAlaLeuA 1723
            |||:::|||   ||||||      ||||||         :::      ||||||::::
Db   9322 TCGATGTGCTGACGGGGCTCGGAAGGCCCCGCGTCGGAGGCCAACAATGCCTTGCGGTCG 9263

Qy   1723 snTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeuProValL 1743
          ::   |||   |||   :::                  ::::::   |||   |||
Db   9262 ACCTTGCCGTTGGGCGTCAG------------CGGCAGCGCCTCCAGCACCA-CCA---- 9220

Qy   1743 euValProProThr---ProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuP 1762
               |||||||||   |||   ||||||||||||:::      |||            |
Db   9219 -----CCACCGACGGCACCATGTACTCCGGCAACTTCTGCCTCACGT---------ACTC 9174

Qy   1762 roThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGlyP 1782
          ||:::||||||   |||                     :::|||      |||      |
Db   9173 CTTCAGCGCCCCTCCCTCTCCTCCCGGCGCCACCACGTACGCCT---CCACCCGCTTCTC 9117

Qy   1782 roThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArgAspA 1802
          ||      |||      |||
Db   9116 CTGCCCCTCTCCCCTCACCACCGCTGCCG------------------------------- 9088

Qy   1802 rgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThrValG 1822
                                                |||   |||
Db   9087 --------------------------------------CTTCCTTCACCTGCGCGTGCTG 9066

Qy   1822 luHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySerSerG 1842
               ||||||         |||   |||:::      ||||||:::         |||
Db   9065 CTTCAGCGCCTCCTCCACCTCCCCCAACTCAA------TCCGGTACCCCCTCACCTTCAC 9012

Qy   1842 lyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHisSerP 1862
                        :::         ||||||:::|||      |||            |
Db   9011 CTGCTCGTCCCGACGCCCCAGAAACTCCAGCAGCCCACCCTCCTGCCACCTCACCTCGTC 8952

Qy   1862 roIleSer------------------ProArgThrGlnAspAlaLeuGlnGlnArgProS 1876
          ||   |||                  |||:::|||      |||            ||||
Db   8951 CCCCGTCCGGTACACCCGCTCCCCCTCCCCAAACGGTGACGGCACGAAACGCTCCGCCGT 8892

Qy   1876 erValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProT 1896
          ||         :::                                 |||      |||
Db   8891 CCACTCCGGCCTCC---------------------CCACGTACCCCTCCGCCAGCCCCTC 8853

Qy   1896 hrValLeuArgSerThrSerThrSerSerPro------------------ValArgProA 1910
            :::   |||:::|||      |||   |||                  :::|||||||
Db   8852 TCCTCCCACGTACAACTCTCCCCTCACCCCCACCGGCACCGGACGCATTCCTTCGTCCAG 8793

Qy   1910 laAlaThr---------PheProProAlaThrHisCys---------------ProLeuG 1922
          ||   |||         ||||||||||||:::                     |||
Db   8792 CACGTACACCCGCGTATTTCCCACCGGCTTCCCAATGGGCACCGTCAACCCCACCT---- 8737

Qy   1922 lyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGluA 1942
                                 |||   |||            |||      |||   :
Db   8736 -----------------------CCTCCTCTGCGACGCGGTGCACTGTCGTGAAAGTCGT 8700

Qy   1942 laProArgValAlaArg-ProGluArgProArgAlaAspThrGlyHisAlaPheLeuAla 1961
          ::|||      |||||| |||            |||   |||   |||:::   |||
Db   8699 CCCCTCTGTCGGCCCGTACCCGTTGAGGAGGATGCCCCCACTCCTCACTCT---CTCCCT 8643

Qy   1962 LysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGluPro 1981
             ||||||   |||      :::            :::|||:::
Db   8642 CACCCGCCCCACCGACAGCACGTCTCCTC-CCGCCAGCACCTGCCTCACTCCACTCAGCG 8584

Qy   1982 ArgProLeuValProProValSerGlyHisAlaThrIle---AlaArgThrProAlaLys 2000
             |||         |||   |||            :::   :::   |||
Db   8583 CTTCCGGCTGGTACGCCTGCATCTGCTCGAAGAGGGCTGCCGTCAGCCACAGCGACGTCA 8524

Qy   2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAla---SerAlaSer 2019
             |||   |||      ::::::|||   ||||||   |||   |||         :::
Db   8523 CCCCTGCCTCCTTCAGCCTCCGCCCCAGCGCCTCCAACTCCAGCGGCCCTTCCGGGTACA 8464

Qy   2020 AspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArg 2039
             |||         :::      ::::::|||
Db   8463 CCACCAGCTTCCCCCCGTGCAGCAGCGCGCCCC--------------------------- 8431

Qy   2040 SerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProVal 2059
                                                 :::   |||   |||
Db   8430 ---------------------------------------ATACCTCCAACGTCGACGCGT 8410

Qy   2060 SerSerProSerLeu-ThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLy 2079
             |||||||||                   |||      ||||||         :::
Db   8409 CGAAGCCCGTCGGCGCCAA------------TTGCAGCCACACCTCCTCCGGCCCGAAGT 8362

Qy   2079 sSerHisLeuGluGlyGluLeuArgPro---------------LysGlnProGlyProVa 2094
              |||         :::   :::|||               :::         |||::
Db   8361 CCGCCACGCCGCCGCCCAACACCAGCCGCGTTACCGCACGCTGCGGCACCGCCACCCCCT 8302

Qy   2094 lLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnPr 2114
          :::::::         :::            |||:::|||   :::      |||
Db   8301 TCGGATTCCCCGTGCTTCCCGACGTAAACATCACGTACGCCAGGTGGCCTCCTCCTAC-- 8244

Qy   2114 oSerSerSerProLeuLeuGlnThrAlaProGly---------------------ValLy 2127
           |||   |||   :::            ||||||                     :::
Db   8243 -TCGGGCTCCCGGGTTTGTCTCTGGCTGCCTGGCAATCTGCGTGTCCCACTCCGTATCCA 8185

Qy   2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
              |||   ::::::   |||            :::   :::::::::   :::
Db   8184 CGCACACCACCAACTCGCTACCCGACGCCACTTCGTCCGCCAGCTTCTCCTGCGCCACCA 8125

Qy   2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
              ||||||   |||         |||:::         ||||||
Db   8124 GCAGCGCCACCC-CCGCCTCCCGCTTCATCCACCCCAGCCGCTCCA------------GC 8078

Qy   2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
           |||:::|||   |||            |||   |||||||||:::
Db   8077 GGGTAGCTCGCGTCCAGCGGCACGTACACGCCCCCCGCCTTCAGAA-------------- 8032

Qy   2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSer-- 2206
```

```
                           :::|||||||||      |||                  |||:::
Db       8031 -------------TCCCCAGCGCGCTCACCACCCACTCCAACGAGCGCTCCGCGCACAAT 7985

Qy       2207 -------------------------ProGluProAsnLysThrSerValLeuGlyGlyGl 2218
                                       |||   |||:::               ||||||
Db       7984 CCCACTCGCACCTCGGGGCCCACTCCCATTCCCCTCA---------------GGTGGTGC 7940

Qy       2218 yGluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerAr 2238
                     |||:::                     |||   :::   |||   :::|||
Db       7939 GCCAACTGGTTTG------------------CTCGGCGATTCAACGCCTCGTAACTCACC 7898

Qy       2238 gSerAlaVal---TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMe 2257
               :::|||:::      ||||||                           |||   |||
Db       7897 TGCTGCCCTGCCCACTCCACTG---------------------------CCACCGCGCCC 7865

Qy       2257 tGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheSer----------- 2273
               ||||||      |||         :::   |||||||||      |||
Db       7864 GGGGTCCTCTCCACCTGCGCTTCAAACAGCTCCGCCAGCGAC-GCCTCCCTCGGGTACTC 7806

Qy       2274 ----LysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
                  :::|||         :::|||         :::            :::||||||
Db       7805 CGCTCGCCTCCCGCTCCACTCCTCCACAAGGACGCGCCGCTCCTCCACCGACATCAACGA 7746

Qy       2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsn-------------------- 2305
                     :::         |||:::||||||      :::
Db       7745 CAGCTCAGACAAGGTCCGGCGCTCCGAACCACACAGTTCGCGCAGGAGCTGCCGGGCCTG 7686

Qy       2305 ------------------------------------------------------------ 2305

Db       7685 CTCGACCATCCGCTTCGCCGCCGATGCATCAAAGCGCCTGGCGTCGTAGGTCAGGCGCAG 7626

Qy       2306 ----IleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLe 2324
                     |||:::|||||||||   |||               :::||||||
Db       7625 CTGCCATTCACGCCCCGGGACGGCCATCATC------------GTCACCGGATGT----- 7583

Qy       2324 uMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAl 2344
                           |||   :::            :::   |||            :::
Db       7582 ----------GAGTCCACCTCGACGATGTACGGGTCGCGAACCACCAGCTCGTCCGACAT 7533

Qy       2344 aIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSe 2364
               ::::::|||   |||      |||   :::            ::::::      |||||
Db       7532 CGTGGTGAGGCCCGCATCCCGGGGGTAGTTC---------TCGAAGACGACCAGGCTCTC 7482

Qy       2364 rAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeu-ProAlaAlaMetProIleT 2384
              |   ||||||      |||               ||||||       |||   |||
Db       7481 GAACAACGCG---GGGCCA------GGGGGAACCTCGCTCCAGCGGCGCACCTCCA---- 7435

Qy       2384 hrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaL 2404
                         |||   :::   ||||||   :::   ||||||            |||:
Db       7434 --------CCAGGGGCGACGTCTCATACTGCCTCGCCTCCTGGA------------GCCA 7395

Qy       2404 ysValSerGlyArgProSerSerArg------LysAlaLysSerProAlaProGlyLeuA 2422
              ::::::::      ||||||:::|||      :::      :::   |||      ||||
Db       7394 GGCTTGCAGCC---CCTTCAGCCAGGGGCCCACGGGGCTCGACGCGGGCAGGCGCACTCG 7338

Qy       2422 laSerGlyAspArg------------ProProSerValSerSerValHisSerGluGlyA 2438
              ||   |||   |||            |||||||||               :::   |||
Db       7337 CACGGGGATGGCGTTGATGAACAAGCCCACCATCGCCTCGATGCCGGCAAGCTCCGGGGG 7278

Qy       2438 spCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaG 2458
                |||   ||||||:::      |||   |||            ||||||
Db       7277 CCTGCCGGAGACGGTCAACCCGAACAGGACGT------------CGTCCTCTCCGGTATA 7230

Qy       2458 lySerThrProPheProTyrAsnProLeuIleMetArgLeu------------GlnAlaG 2474
                :::|||               ||||||                           ||||
Db       7229 GCGCCACAGCAACAGCGCCCAGGCCGCTTGCAACAGCGTGCCTCCCGTGACGCCCTGCTG 7170

Qy       2474 lyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAlaGlyP 2494
              ||            |||   ||||||      |||      ||||||            |
Db       7169 GCGCGAGAAGGCCTCCAGTGCCGCCGTCTCCTCCGCGCTCATCCGGAGCAGCACTTCACC 7110

Qy       2494 ro---------HisHisAla 2497
              ||         |||||||||
```

```
Db       7109 CAGGAGCGTCACATCACGCA 7090

RESULT 7
US-09-902-540-9597/c
; Sequence 9597, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9597
;   LENGTH: 14462
;   TYPE: DNA
;   ORGANISM: Myxococcus xanthus
US-09-902-540-9597

Alignment Scores:
Pred. No.:              7.98e-23        Length:         14462
Score:                  698.00          Matches:        652
Percent Similarity:     32.50%          Conservative:   301
Best Local Similarity:  22.24%          Mismatches:     1028
Query Match:            5.28%           Indels:         964
DB:                     4               Gaps:           143

US-09-522-753-5 (1-2517) x US-09-902-540-9597 (1-14462)

Qy         44 HisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGlnProGlnArg 63
              |||   ||||||                  |||   |||:::         |||
Db      14437 CACCCCAGCCGCTCCAGCGGGTAGCTCGCGTCCAGCGGCACGTACACGCCCCCCGCCTTC 14378

Qy         64 ArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSer-GlnGluLeuHi 83
              |||   ||||||   |||:::               |||||||||||| :::   :::
Db      14377 AGAATCCCCAGCGCGCTCACCACCCAC------TCCAACGAGCGCTCCGCGCACAATCCC 14324

Qy         83 sLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIl 103
               |||   |||
Db      14323 ACTCGCACCTCGGGGC-------------------------------------------- 14308

Qy        103 eGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgProSerProLe 123
                                                               |||   |||
Db      14307 -------------------------------------------------CCACTCCCATT 14297

Qy        123 uLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThrGl 143
                  :::   |||   |||   |||            :::      |||:::
Db      14296 CCCCTCAGGTGGTGCGCCAACTGGTTTGCTCGGCGATTCAACGCCTCGTAACTCACCTGC 14237

Qy        143 yLysLeuGluProValSerProProSerProProHisThrAspProGluLeuGluLeuVa 163
                  |||            ||||||   |||      :::   |||   |||:::
Db      14236 TGCCCTGCCCACTCCACTGCCACCGCGCCCGGGTCCTCTCCACCTGCGCTTCAAACAGC 14177

Qy        163 lProPro-ArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGlu- 182
               |||||| |||                           :::   ||||||   |||
Db      14176 TCCGCCAGCGA------------------------CGCCTCCCTCGGGTACTCCGCTCGC 14141

Qy        183 -----IleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuG 201
                         ::::::   |||            |||   :::   |||:::||||||
Db      14140 CTCCCGCTCCACTCCTCCACCAGGACGCGCCGCTCCTCCACCGACATCAACGACAGCTCG 14081

Qy        201 luGluGluAlaAlaLysProProGluProGluLysProValSerPro------------- 216
                :::   |||:::            |||      |||   :::|||
Db      14080 CCCAGGC-GCGTCTCCGGATTCGCGACCGCGGCCTCCAGCAGCACCGACATCTGCCTCAG 14022

Qy        217 --------ProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluA 234
                      ||||||      :::   :::|||||||||:::   :::
Db      14021 CAGCCGCTCCGCCGTCTTC-GCCTCGAACAGGTCGCTGTTGTAATTG------------- 13976
```

```
Qy          234 snArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluL 254
                           |||   |||      |||   |||:::            ||||||||||
Db        13975 -----ACTGACGCGGCGGCACC-TTCCGGCAGCTCGACGACGAACATCGACAGGTCGAAC 13922

Qy          254 euProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnG 274
                ||         |||                                    :::::::::
Db        13921 TTCGACGTCTGAATCCG---------------------------------CGCGTCCATC 13895

Qy          274 lnAlaMetArgLysLysLeuIle---LeuTyrPheLysArgArgAsnHisAlaArgLysG 293
                        |||:::   |||      |||:::|||      |||         :::   |
Db        13894 ACCGTCAGCGACAATCCCTGTCGCAGCTTCACTTCGGGCACAGGCGTGTTCTGAAGTGTC 13835

Qy          293 ln------TrpLysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluL 311
                ||         ||||||:::         :::            ::::::   |||   :
Db        13834 AGGGAGACTTGAAACAACGGGCTGCGACTCAGGTCCCGCTCCGGCTGCAATGCCTCCACC 13775

Qy          311 ysLysValGluArgIleGlu------------------------------AsnAsnProA 321
                ::   :::||||||                                       :::   |
Db        13774 AGCTTCTCGAACGGGACGTCCTGGTGCGCGTAGGCGCCCAGGGTCACCTCACGCACCTGC 13715

Qy          321 rgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleA 341
                ||:::         |||   |||:::      |||         |||         |||
Db        13714 GCCAACATGTCCCGGAAGGAAAGCTCGGGACGTACGCGAGCCCTCAGCACCAGGGTATTG 13655

Qy          341 rgLysGlnArgGluLeuGlnGluArgMet---------GlnSerArgValGlyGlnArgG 358
                  :::      :::         |||:::            :::|||   |||:::
Db        13654 GCGAAGTTGCCGATGAGCCCCTCCGTCTCGGCCTGCGTGCGACCCGCGATGGGGAAGCCC 13595

Qy          358 lySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspG 378
                  :::   :::   :::               :::|||   :::
Db        13594 ACGCTGATGTCGTCCTGCCCCGAGAGCCGCCACAACACCGTCTG---------------- 13551

Qy          378 lyLeuSerGluGlnGluAsnLeuGluLysGln------------MetArgGlnLeuAlaV 394
                  :::      |||::::::   |||::::::               |||
Db        13550 --ATACGCCGCCAGCAGCACCATGAACGGCGTGGCCCCTTCTCGCCGCGGTCGCCTCC 13493

Qy          394 alIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnG 414
                  :::||||||                     ||||||::::::
Db        13492 ACTTGCCGCCA---------------------CAGCGCCTTCGG---------------- 13470

Qy          414 lyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerG 434

Db        13470 ------------------------------------------------------------ 13470

Qy          434 luGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuI 454
                                                      |||
Db        13469 -----------------------------------CCACAT------------------- 13464

Qy          454 leAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrL 474

Db        13463 --------------------------------------------------------CTGC 13460

Qy          474 ysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerG 494
                     :::         |||   |||   |||   :::         ||||||:::   |
Db        13459 CGCTGGATTCCACCTCGAAAGGTCTGCGCCGCGGGGCGAGGCCTGTCCGTGGGAGGTCC 13400

Qy          494 lnGlnGlnGlnGlnGlnGlnGln------------------------------------- 502
                ||:::   ::::::   :::|||
Db        13399 AGGAATCGCGGCGCGCCCGCCAGCACCTCGCGCCAGTAGCGAACCTGCGATTCCAACACC 13340

Qy          503 -----------------------------------------------------GlnGlnG 505
                                                                     |||   :
Db        13339 TCGTCCTTCAACCACTCGCGCTGCCAGAGCGCGTAGTCCGCGTACTGCACGGGCAGCTCC 13280

Qy          505 lnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluL 525
                ::|||      :::||||||:::   |||            |||:::   :::      |
Db        13279 GGCAATGGTGACGGCAGCCCTTG---CGAGAACGCGTCGTAGAGCGCCGCCAACTCCCGA 13223

Qy          525 ysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsnAspLysGluAspL 545
                ||:::                                 ::::::         ||||||:
Db        13222 AGCAGCACGCTCGCGGACCAACCGTCCGAGACAATGTGATGCAACAGCAGCACGAGGACG 13163

Qy          545 euLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValA 565
                ::         :::|||      |||:::|||
Db        13162 TGCG-CCTGCTCGGACTGCTTCAACAGCAGGG---------------------------- 13132

Qy          565 laSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyArgIleThrArgS 585
                        |||         |||:::::::::   |||   :::||||||:::   |||:
Db        13131 --------GGCGCAGCAGCGGCCCACGCGCGAGCTCGAATGCGCGGTCGACTTCCGCGCG 13080

Qy          585 erMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGlnSerAlaGluLeuA 605
                ::            |||   |||:::   :::      |||      :::   |||
Db        13079 CA---CCCGGTGCCGCACC-TCGGACTCATCCTGGGCGCCCAGCTCCACCACGGAGAGGT 13024

Qy          605 laSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysL 625
                  :::   :::      |||:::   |||||||||               :::   |||
Db        13023 GGACCGGCGAT---GGCGAAGCGATGCGCTGGACAGGTGCTCCGCCTTCATCCGGGAAG- 12968

Qy          625 ysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysT 645
                                                |||:::   |||            :::|
Db        12967 --------------------------------GCCGTACGCAGGACTTCATGACGGCGCA 12940

Qy          645 hrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAsp---- 663
                ||   |||:::                        |||   |||         :::
Db        12939 CCAGCTCCGAG------------------------AAGGCTCGCTCCAAGGCAGAAACAT 12904

Qy          664 --------GluIleLeuGlnGln-HisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                                 :::    |||   ::::::      :::   |||   |||
Db        12903 CCAGGGGGCCCTCCACCCGCAGCACATACGGTGCGTTGTACTCGGTACCGCCCGGCTTCA 12844

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAla---PheProProVal 699
                                     |||||||||         ||||||      ||||||
Db        12843 GCCGATCCAGGAACCACAGTCGCTGCTGCGCGAATGACAGCGGCACCCGGTCCCCCCGAG 12784

Qy          700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
                      |||            ||||||
Db        12783 GCATCGGACGCAGCGCGAGCAGCGTCTGACTTC--------------------------- 12751

Qy          720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluVal------------ProArgGlyGlu 735
                |||:::   |||      |||                              ||||||
Db        12750 GAGTCGCCTGCTCCACCTGCTGGGCAAAGGACCGCAAGGTGGGCGCCTCGAAGAGGGCGC 12691

Qy          736 CysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHis 755
                   :::||||||                                    |||   ||||||
Db        12690 GTAGCGGGACCTCTACGC------------------------------CCAGGACCGCAC 12661

Qy          756 ThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAla 775
                      |||                        |||   |||         :::
Db        12660 GCACACGCG------------------------CCACCACCTGCGTGGCCAGTAGCGAAT 12625

Qy          776 AspGlyPro---------------ProProGlyProProThrProProArgArgThrSer 790
                |||   |||               |||||||||||||||   ||||||   |||
Db        12624 GACCACCCGACGCGAAGAAGTCGTCCTCCCGGCCCACCCGCTCCACCCGCAAGA------ 12571

Qy          791 ArgAlaProIleGlu---------------------------ProThrProAlaSerGlu 801
                |||   |||   :::                           |||   |||
Db        12570 CGTGGGCCCAGACAGCGGCCAACGACTCTTCCATGGGGGAAGCCGGCGCCTCATAGGCGG 12511

Qy          802 AlaThrGly---AlaProThr---------------------------ProPro------ 809
                      |||   |||||||||                           ||||||
Db        12510 GAGCCGGGCCACGCGCCGACACGTTCGGCACGGGCAGCGCGCCAGTCCACCTTGCCAT 12451

Qy          810 ---------------------ProAlaProProSerProSerAlaPro------------ 818
                                     ||||||||||||:::|||:::||||||
Db        12450 TGGGCGTCAGCGGCAGGGCGTCCAGCACCACCAGCGCCGACGGCACCATGTGCCCAGGAA 12391

Qy          818 ------------------------------------------------------------ 818

Db        12390 GCCGCTGCTGGAGGAAGCCCCTGAGGGCCGCGATGTCCGGAGCCGCCCTGGCCGTACGAA 12331

Qy          819 ---------ProProVal---------------------------------ValProLys 824
                         ||||||                                       |||:::
Db        12330 CGGCGTAGCCCACCAGCCGATTGTCTCCAGGCACCTCTTCTCGCGCGACCACCACCGCGT 12271

Qy          825 GluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLys 844
```

```
                    :::   :::         |||   ||||||:::      |||      :::
Db      12270 CCTGGACGTCGGGATGCGCTCTGAGCACCGCCGCCACTTCT-CCGGGCTCCACCCGGAAC 12212

Qy        845 ProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSer 864
              |||                           |||   |||:::      |||:::   |||
Db      12211 CCT-----------------CTCACCTTCACCTGGAAGTCCGTCCGCCCGATGAACTCC 12170

Qy        865 Glu------CysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAla 882
                                               |||   :::         |||   |||
Db      12169 AGTCGCCCATCCCCCAACCACCGCACCTTGTCCCCCGTGCGGTACATCCTCGCTCCAGCC 12110

Qy        883 AlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArg 902
              :::               |||      ::::::                  |||||||||
Db      12109 TCCATGCTGAACGGGTCCGGCACGTACCGCTCCGCCGTC---------AACTCCGGGCGC 12059

Qy        903 AlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSer 922
                 |||      :::   :::   ||||||      ::::::|||      |||   :::
Db      12058 CATACGTACCCTCGCGTGACCTGCGCCCCTCCCACGAACAACTCTCCCCCACGCCCACC 11999

Qy        923 AlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer-ProArgPr 942
                                         |||         ||||||   |||    |||
Db      11998 ---------------------------GGCACAGGACGCAGGCTCCGGTCCAGCACGTAA 11966

Qy        942 oSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------Pr 960
                  :::      |||:::      |||         |||:::|||   :::      ||
Db      11965 ACCCGTGTGTTGCCCATCGGCCGTCCCA---------GCGGCACCGTCGCGGGTGACTCC 11915

Qy        960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal----- 978
              |||||||:::                  |||         ||||||
Db      11914 GCTCGACGTG---------------GCAGCTCCACTCGCCCCGCCAGCACACCCACCGTC 11870

Qy        979 ----------------------------------ThrLysValHisGluProProArgGl 987
                                                |||:::         ||||||
Db      11869 GTCTCCGTCGGGCCGTAGTGGTTGTGCACCTCGCACCCGGGCGCCAGCGCCGCCACCTGC 11810

Qy        987 uAspAlaAlaProThrLysProAlaProProAlaProProPro----------------- 1001
                  |||      |||:::|||   |||      |||||||||
Db      11809 TCCAGCAGGCTCCACGAGGCCG---CCTCTCCTCCCACCACCAGCCGCTTCGTCGGCAAC 11753

Qy       1002 -------------------------------ProGlnAsnLeuGln------ProGluSe 1009
                                             |||:::      :::      |||   ::
Db      11752 ACCTGCCCGGGCGCGGCGCCACTCATCAGCGCCGCGAGGTGCGAGGGCACCACCTTCAAA 11693

Qy       1009 rAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAlaProPr 1029
              :      |||         ||||||:::|||            |||            ||
Db      11692 CACTCCACCCCATGCCGCTGGAAGTACTCCTCCACTCCCGCCGCGCTGCTTGCCCTCTCC 11633

Qy       1029 oAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTr 1049
              ||||         |||   |||                  |||         ||||||::
Db      11632 TGCGTCAGCACATGCA---GCAGTCCCCCCGTCCACAGCGCCGCGTACAGCACCGTGTTT 11576

Qy       1049 pThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAl 1069
              :            |||   |||   ||||||            ::::::|||
Db      11575 C------------CCAGGTCCGCCGCCACCGTCGACACCAGCGCGAAGCTCGTGCACTCC 11528

Qy       1069 aProAspProSerAlaPheSerTyrAlaProProGly----------------------- 1081
               |||   |||:::|||   |||      |||
Db      11527 GCCAACCCCAGCCGCGCATTCGCCGCTTCCACGTAGTGGGCAAGCTGTCCGTGTGTGACC 11468

Qy       1081 ------------------------------------------------------------ 1081

Db      11467 GCCACGCCCTTGGCCGCCCCGTGCTTCCGGATGTGAACAGCACGTACGCCAGTTGCTCG 11408

Qy       1082 ----HisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg----- 1098
                     |||||||||      |||         |||            ||||||
Db      11407 GGTTGCACCTCTACCCGCGGTGCTTCCGTCGGCTGCGCCTCGAGCCGCGCCGCGTCTCCA 11348

Qy       1099 -------ProPro-------------------------------------ThrIleSe 1103
                     ||||||                                             ||
Db      11347 TCCATCTCCACCCGGCGCGCGGACAGTCCCTCGAAAGTCGCCGCATGCCGCGCCTGCGTC 11288

Qy       1103 rAsnProPro------ProLeuIleSerSerAlaLysHisProSerValLeuGluArgGl 1121
              |   ||||||      ||||||:::      |||:::         :::|||
```

```
Db      11287 ACCACCACCGGCGCGGCCACTTCTTCCACCAGCGCGCGCAGGCGGGATTCTG-------- 11236

Qy       1121 nIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAl 1141
                        :::::::::|||   :::   :::      :::|||   ||||||
Db      11235 ----------GTTGCGAGGGGTCCAGCGGCACGTACGCGCCTCCCGCCTTCAGAATCCCC 11186

Qy       1141 aLysAlaPro---------------ValGlyProValThrMetGlyLeuProLeuProMe 1156
                  :::|||                     |||   |||      |||   ||||||
Db      11185 AGCATCGCCACCACCCACTCCACGCCCCGCTCCAGGCACAGCGCCACTGGCACTTCCGCG 11126

Qy       1156 tAspProLysLysLeuAlaProPheSerGlyValLysGlnGlu---------GlnLeuSe 1173
                  |||            ||||||
Db      11125 CCCACCC------------CCATTCCCCTCAACTGGTGCGCCACCTGGTTGGCCTGCGCC 11078

Qy       1173 rProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSe 1193
               |||   |||            |||               |||   |||      |||||
Db      11077 TCCAACTGGCGGTACGTCAGCTCCCGCTCCTCGCACCGCACCGCGGGCGCCTCTGGCGTC 11018

Qy       1193 rValLeuArgGly-------ThrAlaLeuGlySerValProGlyGlySerIleThrLysG 1211
              |   |||||||||       |||   :::   |||:::|||      |||:::   :::
Db      11017 CGCTCTGCGTGGTGCTCGACACACCGGTGGACAGCCTCCCACCGGAGAGCCTGGGCCGAG 10958

Qy       1211 lyIlePro------------------SerThrArgValProSerAspSerAlaIleThrT 1225
                :::|||                  :::   |||:::                     |
Db      10957 CCGTTCCATTCCACCAGCACCTGATGACGCTCCGCATCGGTCAGCAG---------CGGT 10907

Qy       1225 yrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrA 1245
              ||   |||            |||         |||            |||
Db      10906 ATCTGGGAGAGGCGCGATTCGGGCCGCGTGGTGAC------------GGCCTTCAGCACC 10859

Qy       1245 rgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAsp---------- 1261
                :::      |||      |||   |||      |||||||||   :::
Db      10858 TGCTCCAGGTAGAACACCATCCGCTCCGC------CGTGGCCGCCGTGAAGAGGTCGGTG 10805

Qy       1262 --------------------------------SerLeuProLysGlyHisValIleTyrG 1271
                                              |||||||||:::               |
Db      10804 CTGAACTGGAGCAAGCCCTCGAAACCCCGTGGAGTCTCCCCCAGGCTCAACGAAAGCTCG 10745

Qy       1271 luGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerL 1291
              ||      :::                           :::   |||   |||      :
Db      10744 AAGCGAGCCGC---------------------------CTCCTTGTCAATCAACCCCATC 10712

Qy       1291 ysGluAspGly---ArgSerSerSerGlyProProHis---------------------- 1302
              ::      |||   |||               ||||||
Db      10711 GGCTTCACGGTCAGCGAGGGCAGCGACAACGGCCCCACCGGGGCGTTCTGCAACACGAAG 10652

Qy       1303 -----------GluThrAlaAlaPro------------------LysArgThrTyrAspM 1313
                         |||   |||||||||                  :::      ::::::
Db      10651 AGGACCTGGAAGAGCGCGCTGCGCCCCAGGTTCCGCTCGGGTTGCAGATGCTCCACGAGC 10592

Qy       1313 etMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyA 1333
              :::::||||||:::|||                                 ||||||
Db      10591 TTCTCGAAGGGAAGGTC---------------------------------CTGATG---- 10569

Qy       1333 rgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHis---------H 1350
                   :::   |||:::|||   :::      ||||||::::::|||
Db      10568 --CTCGTACGCCCCCAGCGT---GCGCTCACGCACCTGCGCCAGCAGTTCGCGGAAGGTG 10514

Qy       1350 isIleArg---------GlySerIleThrGlnGlyIlePro------------------- 1360
                :::|||         |||:::      ||||||:::
Db      10513 GCATGAGGCTGGACCTGGGCGCGGAGAACCAGGGTGTTGACGAAGAAGCCGATGAGCCCC 10454

Qy       1361 --ArgSerTyrValGluAlaGlnGluAsp---------------------TyrLeuArgA 1373
                |||   :::|||   |||      |||                        |||   |
Db      10453 TCCGTCTCTTCGTACCGGCGACCCGCGATGGGCGAGCCCACCAGCACGTCCTCCTGCCCC 10394

Qy       1373 rgGluAlaLysLeuLeuLysArgGluGlyThrPro------------------------P 1385
              ||   |||:::      :::   ||||||   |||                        |
Db      10393 GAGTAGCGCGACAGCAGGAGCTGGAAGGCCG-CCAGCAGCACCATGAAGGGCGTGGCTCC 10335

Qy       1385 roPro------ProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuG 1403
              |||||      |||   |||||||||      :::   |||      |||
Db      10334 CTCCGCATGGGCCAGTGCCGTCACGGCGTCGCTCATCGGCCGTGACAACCGCACCG---- 10279
```

```
Qy     1403 lyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGlyA 1423
                             |||   |||   |||      ||||||         |||
Db    10278 --------------GCACCCATGCACCGCGGT------GCGACAACGTCGCCGGCCTGGG 10239

Qy     1423 rgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuProLeu---- 1441
                             |||   ||||||   |||   |||
Db    10238 CTTGTCGGTGGGTAGCTCCAGGTGAGGAGGCGCGCCAGACAGCTGCTGCCGCCACCAGTC 10179

Qy     1442 --AlaProArgProLeuLysGluGlySer---IleThrGlnGlyThrProLeuLysTyrA 1460
              |||      |||   :::      |||   |||         ||||||
Db    10178 GAGCTGCGTCTCCAGCACGGCGCCGCTCAGCCATTGCCGCTGCCACACCGCGTAGTCCGC 10119

Qy     1460 spThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerP 1480
              |||      :::         :::::::::      :::|||         |||   |
Db    10118 GTACTGCACGGGCAGCTCCGGCAACCCGGGAGACCGTCCTTCGA---------GGAAGGC 10068

Qy     1480 roGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuG 1500
            ||   |||      ||||||      |||                  |||   :::
Db    10067 CGCATAGAGCGCGCCCACCTCGCGGACCA---------GCACGCCCAGCGACCAGCCATC 10017

Qy     1500 luArgAlaCysTyrGluGluSerLeu------LysSerArgProGlyThrAlaSerSerS 1518
              |||   |||                     ::::::||||||   ||||||:::::::
Db    10016 CGAGACGATGTGATGCAGCACCAGCACCAGGGCGTGCACGTCCGGCGACAGCCGCGACAG 9957

Qy     1518 er------------GlyGly------------------------------------SerI 1522
            ::            ||||||                                    |||
Db     9956 CCCCACGCGCACCAGGGGGCCGCGCTCCAGATGAAACGGCCGCAGGGCTGTTCGAGTCGC 9897

Qy     1522 leAlaArgGlyAlaProValIleValPro------------------------------G 1532
              |||||||||               |||
Db     9896 ACGCGCGCGGAATTCGGCCTCGCGCTCCACTTCCGGCAGCGTGCTCAAGTCCTCGAAGGC 9837

Qy     1532 luLeuGlyLysPro---------------------------------------ArgGlnS 1539
                 |||:::|||                                       ||||||:
Db     9836 GAGCTGGACGTCCATGGCCGGATGGATGACCTGGACAGGCTGGCCCTCGTGGAGGCAAA 9777

Qy     1539 erProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerP 1559
            ::                  ||||||||||||   :::            |||   :::
Db     9776 CGTCGTGCGCA---ACGACTCATGGCGCGCCACC-AGCGCGGAGAACGCGCGCTGCAACG 9721

Qy     1559 roValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysA 1579
              :::|||      ||||||:::   |||      :::|||   :::   :::      |
Db     9720 CGCTCACATCGAGCGAGCCCTCGATCCGGAG-GACGATGGGGACGTTGTAGCTGGCACTG 9662

Qy     1579 laSerGlnAspArgLysLeuThrSerThrProArgGluIle---------------AlaL 1594
            ||      :::      |||         |||      :::                  :
Db     9661 CCCGGATGGAG------CTGGTCGATGAACCAGAGCCGCTGCTGGGCGAACGACACGGGC 9608

Qy     1594 ysSerProHisSerThrValProGluHisHisProHis---------------------- 1606
            ::      |||:::      |||            |||
Db     9607 AATGCGCCCATGCGAGGTACCCGGGTCAGCGCTGGCACGCTCGGCGCCGCCGCTTCCCGC 9548

Qy     1607 -----------------ProIleSerProTyrGluHis---------------------- 1613
                                      |||:::||||||
Db     9547 GCCGACTCCTCGATTCGCAGCGCGAGCCCTTCGAGCACAGGTGACGCGAAGAGCGTGCGC 9488

Qy     1614 -----LeuLeuArgGlyValSerGlyValAspLeuTyr---------------ArgSerH 1627
                    |||         ||||||   |||:::                  :::
Db     9487 AGGGGCAACTCCACGCCGAAAGCGGAGCGGATGTGCGCGATGAGCTGCGTGGCCAGCAAC 9428

Qy     1627 isIleProLeuAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaA 1647
              :::|||         :::      |||:::   |||      ||||||||||||
Db     9427 GAGTGCCCGCCCAACTCGAAGAAGTCAGCGTC---CGCGCTCACCCGCTCGATGCC---- 9375

Qy     1647 laAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProP 1667
                             :::|||      |||
Db     9374 -----------------CAGCACTTCACACCA---------------------------- 9360

Qy     1667 roTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleA 1687
                                                  ::::::   |||   ::::::
Db     9359 --------------------------------------GATGATGCTCAGCTTCTGCTCG 9338

Qy     1687 snAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgA 1707
              |||                                                   |||
Db     9337 GTGAC---------------------------------------------CGTACGCGGC 9323

Qy     1707 laAspMetLeuArgGlyLeuSer---ProArg---------GluSerSerLeuAlaLeuA 1723
              |||:::|||   ||||||      ||||||         :::      ||||||::::
Db     9322 TCGATGTGCTGACGGGGCTCGGAAGGCCCCGCGTCGGAGGCCAACAATGCCTTGCGGTCG 9263

Qy     1723 snTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeuProValL 1743
            ::   |||   |||   :::                  ::::::   |||   |||
Db     9262 ACCTTGCCGTTGGGCGTCAG------------CGGCAGCGCCTCCAGCACCA-CCA---- 9220

Qy     1743 euValProProThr---ProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuP 1762
                 |||||||||   |||   ||||||||||||:::      |||            |
Db     9219 -----CCACCGACGGCACCATGTACTCCGGCAACTTCTGCCTCACGT---------ACTC 9174

Qy     1762 roThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGlyP 1782
            ||:::||||||   |||                     :::|||      |||      |
Db     9173 CTTCAGCGCCCCTCCCTCTCCTCCCGGCGCCACCACGTACGCCT---CCACCCGCTTCTC 9117

Qy     1782 roThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArgAspA 1802
            ||      |||      |||
Db     9116 CTGCCCCTCTCCCCTCACCACCGCTGCCG------------------------------- 9088

Qy     1802 rgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThrValG 1822
                                                  |||   |||
Db     9087 --------------------------------------CTTCCTTCACCTGCGCGTGCTG 9066

Qy     1822 luHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySerSerG 1842
                 ||||||         |||   |||:::      ||||||:::         |||
Db     9065 CTTCAGCGCCTCCTCCACCTCCCCCAACTCAA------TCCGGTACCCCCTCACCTTCAC 9012

Qy     1842 lyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHisSerP 1862
                          :::         ||||||:::|||      |||            |
Db     9011 CTGCTCGTCCCGACGCCCCAGAAACTCCAGCAGCCCACCCTCCTGCCACCTCACCTCGTC 8952

Qy     1862 roIleSer------------------ProArgThrGlnAspAlaLeuGlnGlnArgProS 1876
            ||   |||                  |||:::|||      |||            ||||
Db     8951 CCCCGTCCGGTACACCCGCTCCCCCTCCCCAAACGGTGACGGCACGAAACGCTCCGCCGT 8892

Qy     1876 erValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProT 1896
            ||         :::                                 |||      |||
Db     8891 CCACTCCGGCCTCC---------------------CCACGTACCCCTCCGCCAGCCCCTC 8853

Qy     1896 hrValLeuArgSerThrSerThrSerSerPro------------------ValArgProA 1910
              :::   |||:::|||      |||   |||                  :::|||||||
Db     8852 TCCTCCCACGTACAACTCTCCCCTCACCCCCACCGGCACCGGACGCATTCCTTCGTCCAG 8793

Qy     1910 laAlaThr---------PheProProAlaThrHisCys---------------ProLeuG 1922
            ||   |||         ||||||||||||:::                     |||
Db     8792 CACGTACACCCGCGTATTTCCCACCGGCTTCCCAATGGGCACCGTCAACCCCACCT---- 8737

Qy     1922 lyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGluA 1942
                                   |||   |||            |||      |||   :
Db     8736 -----------------------CCTCCTCTGCGACGCGGTGCACTGTCGTGAAAGTCGT 8700

Qy     1942 laProArgValAlaArg-ProGluArgProArgAlaAspThrGlyHisAlaPheLeuAla 1961
            ::|||      |||||| |||            |||   |||   |||:::   |||
Db     8699 CCCCTCTGTCGGCCCGTACCCGTTGAGGAGGATGCCCCCACTCCTCACTCT---CTCCCT 8643

Qy     1962 LysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGluPro 1981
               ||||||   |||      :::            :::|||:::
Db     8642 CACCCGCCCCACCGACAGCACGTCTCCTC-CCGCCAGCACCTGCCTCACTCCACTCAGCG 8584

Qy     1982 ArgProLeuValProProValSerGlyHisAlaThrIle---AlaArgThrProAlaLys 2000
               |||         |||   |||            :::   :::   |||
Db     8583 CTTCCGGCTGGTACGCCTGCATCTGCTCGAAGAGGGCTGCCGTCAGCCACAGCGACGTCA 8524

Qy     2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAla---SerAlaSer 2019
               |||   |||      ::::::|||   ||||||   |||   |||         :::
Db     8523 CCCCTGCCTCCTTCAGCCTCCGCCCCAGCGCCTCCAACTCCAGCGGCCCTTCCGGGTACA 8464

Qy     2020 AspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArg 2039
```

```
                 |||         :::      ::::::|||
Db        8463 CCACCAGCTTCCCCCCGTGCAGCAGCGCGCCCC--------------------------- 8431

Qy        2040 SerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProVal 2059
                                                      :::   |||   |||
Db        8430 ---------------------------------------ATACCTCCAACGTCGACGCGT 8410

Qy        2060 SerSerProSerLeu-ThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLy 2079
                  |||||||||                   |||      ||||||         :::
Db        8409 CGAAGCCCGTCGGCGCCAA------------TTGCAGCCACACCTCCTCCGGCCCGAAGT 8362

Qy        2079 sSerHisLeuGluGlyGluLeuArgPro---------------LysGlnProGlyProVa 2094
                   |||         :::   :::|||               :::         |||::
Db        8361 CCGCCACGCCGCCGCCCAACACCAGCCGCGTTACCGCACGCTGCGGCACCGCCACCCCCT 8302

Qy        2094 lLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnPr 2114
               :::::::         :::            |||:::|||   :::      |||
Db        8301 TCGGATTCCCCGTGCTTCCCGACGTAAACATCACGTACGCCAGGTGGCCTCCTCCTAC-- 8244

Qy        2114 oSerSerSerProLeuLeuGlnThrAlaProGly---------------------ValLy 2127
                |||   |||   :::            ||||||                     :::
Db        8243 -TCGGGCTCCCGGGTTTGTCTCTGGCTGCCTGGCAATCTGCGTGTCCCACTCCGTATCCA 8185

Qy        2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
                   |||   ::::::   |||            :::   :::::::::   :::
Db        8184 CGCACACCACCAACTCGCTACCCGACGCCACTTCGTCCGCCAGCTTCTCCTGCGCCACCA 8125

Qy        2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
                   ||||||   |||         |||:::         ||||||
Db        8124 GCAGCGCCACCC-CCGCCTCCCGCTTCATCCACCCCAGCCGCTCCA------------GC 8078

Qy        2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
                |||:::|||   |||            |||   |||||||||:::
Db        8077 GGGTAGCTCGCGTCCAGCGGCACGTACACGCCCCCCGCCTTCAGAA-------------- 8032

Qy        2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSer-- 2206
                            :::|||||||||      |||                  |||:::
Db        8031 -------------TCCCCAGCGCGCTCACCACCCACTCCAACGAGCGCTCCGCGCACAAT 7985

Qy        2207 -------------------------ProGluProAsnLysThrSerValLeuGlyGlyGl 2218
                                        |||   |||:::               ||||||
Db        7984 CCCACTCGCACCTCGGGGCCCACTCCCATTCCCCTCA---------------GGTGGTGC 7940

Qy        2218 yGluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerAr 2238
                      |||:::                     |||   :::   |||   :::|||
Db        7939 GCCAACTGGTTTG------------------CTCGGCGATTCAACGCCTCGTAACTCACC 7898

Qy        2238 gSerAlaVal---TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMe 2257
                :::|||:::      ||||||                           |||   |||
Db        7897 TGCTGCCCTGCCCACTCCACTG---------------------------CCACCGCGCCC 7865

Qy        2257 tGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheSer----------- 2273
                ||||||      |||         :::   |||||||||      |||
Db        7864 GGGGTCCTCTCCACCTGCGCTTCAAACAGCTCCGCCAGCGAC-GCCTCCCTCGGGTACTC 7806

Qy        2274 ----LysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
                   :::|||         :::|||         :::            :::||||||
Db        7805 CGCTCGCCTCCCGCTCCACTCCTCCACAAGGACGCGCCGCTCCTCCACCGACATCAACGA 7746

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsn-------------------- 2305
                      :::         |||:::||||||      :::
Db        7745 CAGCTCAGACAAGGTCCGGCGCTCCGAACCACACAGTTCGCGCAGGAGCTGCCGGGCCTG 7686

Qy        2305 ------------------------------------------------------------ 2305

Db        7685 CTCGACCATCCGCTTCGCCGCCGATGCATCAAAGCGCCTGGCGTCGTAGGTCAGGCGCAG 7626

Qy        2306 ----IleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLe 2324
                      |||:::|||||||||   |||               :::||||||
Db        7625 CTGCCATTCACGCCCCGGGACGGCCATCATC------------GTCACCGGATGT----- 7583

Qy        2324 uMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAl 2344
                            |||   :::            :::   |||            :::
Db        7582 ----------GAGTCCACCTCGACGATGTACGGGTCGCGAACCACCAGCTCGTCCGACAT 7533

Qy        2344 aIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSe 2364
                ::::::|||   |||      |||   :::            ::::::      |||||
Db        7532 CGTGGTGAGGCCCGCATCCCGGGGGTAGTTC---------TCGAAGACGACCAGGCTCTC 7482

Qy        2364 rAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeu-ProAlaAlaMetProIleT 2384
               |   ||||||      |||               ||||||       |||   |||
Db        7481 GAACAACGCG---GGGCCA------GGGGGAACCTCGCTCCAGCGGCGCACCTCCA---- 7435

Qy        2384 hrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaL 2404
                          |||   :::   ||||||   :::   ||||||            |||:
Db        7434 --------CCAGGGGCGACGTCTCATACTGCCTCGCCTCCTGGA------------GCCA 7395

Qy        2404 ysValSerGlyArgProSerSerArg------LysAlaLysSerProAlaProGlyLeuA 2422
               ::::::::      ||||||:::|||      :::      :::   |||      ||||
Db        7394 GGCTTGCAGCC---CCTTCAGCCAGGGGCCCACGGGGCTCGACGCGGGCAGGCGCACTCG 7338

Qy        2422 laSerGlyAspArg------------ProProSerValSerSerValHisSerGluGlyA 2438
               ||   |||   |||            |||||||||               :::   |||
Db        7337 CACGGGGATGGCGTTGATGAACAAGCCCACCATCGCCTCGATGCCGGCAAGCTCCGGGGG 7278

Qy        2438 spCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaG 2458
                 |||   ||||||:::      |||   |||            ||||||
Db        7277 CCTGCCGGAGACGGTCAACCCGAACAGGACGT------------CGTCCTCTCCGGTATA 7230

Qy        2458 lySerThrProPheProTyrAsnProLeuIleMetArgLeu------------GlnAlaG 2474
                 :::|||               ||||||                           ||||
Db        7229 GCGCCACAGCAACAGCGCCCAGGCCGCTTGCAACAGCGTGCCTCCCGTGACGCCCTGCTG 7170

Qy        2474 lyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAlaGlyP 2494
               ||            |||   ||||||      |||      ||||||            |
Db        7169 GCGCGAGAAGGCCTCCAGTGCCGCCGTCTCCTCCGCGCTCATCCGGAGCAGCACTTCACC 7110

Qy        2494 ro---------HisHisAla 2497
               ||         |||||||||
Db        7109 CAGGAGCGTCACATCACGCA 7090

RESULT 8
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:              1.01e-19        Length:        4411529
Score:                  689.00          Matches:       526
Percent Similarity:     33.44%          Conservative:  207
Best Local Similarity:  24.00%          Mismatches:    872
Query Match:            5.21%           Indels:        599
DB:                     3               Gaps:          91

US-09-522-753-5 (1-2517) x US-09-103-840A-1 (1-4411529)

Qy         546 LeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValAla 565
               ::::::   |||               |||:::      ||||||   |||   :::
```

```
Db     841474 GTTCGTCGGAAGGTGATGGGGCTGGGCGGTGACGGCTTCGACGAATGCGAATGCCTCTGC 841415

Qy        566 SerLysGlyArgLysThrAlaAsnSer------------------GlnGlyArgArg--- 578
              |||:::   |||         ::::::                  :::   ||||||
Db     841414 TCGCGGTGACGGAATCTCGAAGATGCGTCGATTCGTTGCCAGCCGGAGGAACGACGCCCA 841355

Qy        579 ------------------LysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSer 592
                                :::|||   ::::::|||      :::   |||      :::
Db     841354 TACCAGGTTCGGCACTGTGAAGGGGTCGTCAGCCGCGAGCAGTCGATCGAACCAGGGGCG 841295

Qy        593 Glu---------------------GluAlaIleThrProGlnGln--------------- 600
              :::                        |||:::      ||||||
Db     841294 GACGGTTCGGTGATTCGGGTGGTCACCGCGGTGTGCAGCCAGCAGCACGTTGGCGTCGAG 841235

Qy        601 ---------------SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrp 615
                             ||||||   :::                     |||||||||
Db     841234 AAGGAACATTGCCGTTCTGCGCCTATTGGGGGTTACTTCCGGGAGTTCAGTTCCAGGCCC 841175

Qy        616 ThrGluGluGluMetGluThrAla------------------------------------ 623
              :::            :::   |||
Db     841174 TCGTCGAGCACTTCGGACAACGCTCTATTCGAGGTCAGGTCGATGCCTCGCCGCGGACCG 841115

Qy        624 ------------LysLysGlyLeuLeuGluHisGlyArg---------AsnTrpSerAla 636
                          |||   |||   :::      ||||||            |||:::|||
Db     841114 GTGCCGCCGTCGAAAACCGGGACGGTC------GGGCGGGCGCCGCCGACGTGGGCGGCG 841061

Qy        637 Ile--------------------------------------------------------- 637

Db     841060 GCGAACTCCCGCCGAAGGGCGTCCTCGATCACAGCGCCCAGCGATTGACCACGCTCGCGG 841001

Qy        638 AlaArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLys 657
              ||||||   :::   :::   ::::::|||
Db     841000 GCCCGGCGTTTGGCGGCAGCGAGTATTTCA------------------------------ 840971

Qy        658 LysArgGlnAsnLeuAspGluIleLeu------------------------GlnGlnHis 669
                          :::|||   ::::::                        |||||||||
Db     840970 ------TCGGAGATTGACACCGTGGTGCGCATGATGCACACGGTAACGCATCAACAGCAT 840917

Qy        670 LysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAla 689
                 |||:::   |||:::   :::         :::::::::   ::::::   |||
Db     840916 CATCTGCGGTC-GAGCGGCCGAAGCGATGCTGTCAACGCCGCGCGCGGTCGCAGGCCCGC 840858

Qy        690 AlaSerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGly 709
                       :::|||      |||   :::   |||   :::         |||:::|||
Db     840857 CGGCTACATCAAGCCGTTCAGCCCGTTCTCACCGAGCAGCAGCCCGCCGGCGCCGCCGGC 840798

Qy        710 ValSerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsn 729
                 :::|||
Db     840797 GCCGCTGGG--------------------------------------------------- 840789

Qy        730 GluValProArgGlyGluCysSer------------GlyProAlaThrValAsnAsnSer 745
                                |||:::            |||         |||
Db     840788 ------------GCCGGGTGCACCCCCGGATTTCCCGGCGTTACCGCTGTTACCGCCGTT 840741

Qy        746 SerAspThrGluSerIleProSerProHisThrGluAlaAla------------------ 759
              :::         ::::::         |||      ||||||
Db     840740 GCCGCCGTTGCCGCCGTTGCCGATCAGCACGGCGCTGCCGCCGACGCCGCCCTTGCCGCC 840681

Qy        760 ------LysAspThrGlyGlnAsnGly-------ProLysProProAlaThrLeuGlyAl 775
                    :::   |||||||||   |||       |||   ||||||:::
Db     840680 GATGTGGAAAGCACCGGACAAACCGGTGGCGCCGCCGGCGCCGCCGTCACCTCCGAGACC 840621

Qy        775 aAspGlyPro------ProProGlyProProThrProPro------ArgArgThrSerAr 791
               :::   |||      ||||||   ||||||:::||||||               :::
Db     840620 GAACGCTCCGGACTTGCCGCCGGCCCCGCCGTCCCCGCCGGTGGCGGCGCCGGTGGCGGA 840561

Qy        791 gAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProProAl 811
               ||||||      |||   |||      |||   |||      |||      ||||||||
Db     840560 GGCGCCGGCCCCGCCGGCCCCG---CCGGAGCCGACGAGCAAACCCGCGTCGCCGCCGGC 840504

Qy        811 aProProSerProSerAlaPro------------ProProValValProLysGluGluLy 827
              |||||||:::|||      |||            ||||||:::   |||         ::
Db     840503 CCCGCCGGCGCCGCCGTTGCCAGAGTTGCTGAACCCGCCGCTCCCGCCGGCGCCGCCGGA 840444
```

```
Qy        827 sGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLysProProAl 847
              :   :::            |||||||||:::                        |||
Db     840443 GCCGAAGAGCAAGCCTCCGGCGCCGCCGATG------------------------CCGCC 840408

Qy        847 aAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysTh 867
                           ||||||      |||   |||
Db     840407 GACCACGCCGGCGGCGGTCAG-GCTGGAACCGCC-------------------------- 840375

Qy        867 rGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAl 887
                     |||      |||||||||   |||   :::   |||   |||         ||
Db     840374 -------GCTGCCGCCGGCCCCGCCGGAGGCGCCGAGGGAGAGCAAGCCGGCGTCGCCGC 840322

Qy        887 aGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLy 907
              |      |||      :::            ||||||:::      |||      |||
Db     840321 CGGCCCCGCCGGCCCCTCCGGAAACTCCGGCGGCGGTGCTATGGCTGCCGGCGCCGCCGG 840262

Qy        907 sSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluValAs 927
               ::::::||||||   |||   :::   ::::::            |||
Db     840261 TGCCGCCGGCGCCGAACAGCCCGCCGTTGCCGCCCGCCCCGCCGGCGCCGCCGGCGAGGG 840202

Qy        927 pGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThrPr 947
                  :::      |||                  :::   :::|||::::::   |||
Db     840201 TGCTTCCCCCGCCGGCCCCGCCGGTCCCGCCGGCGCCGAACAGCCCGCCGTTCCCACCTG 840142

Qy        947 oThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLy 967
                     ||||||   :::   :::         ||||||         ::::::|||
Db     840141 CCCCGCCGATCCCGCCAGCCCCAGCGGGGCTGCCAAACCCACCGG-GCCCGCCGGCTC-- 840085

Qy        967 sGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArgGl 987
                  |||   |||      ||||||:::                     ||||||
Db     840084 -CCCCGTTG-GCGAACAGCCCGCCGTTA---------------------CCGCCGGCCCC 840048

Qy        987 uAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGlnPr 1007
                        ||||||   |||||||||||||||||||||
Db     840047 GCCATTGCCACCGACCGCGCCGGCACCACCGGCCCCGCCG-------------------- 840008

Qy       1007 oGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAl 1027
                        ||||||      |||   |||:::|||         :::      |||||
Db     840007 ----------GCCCCGGCGGCGCCCAAGAGCGCACCG---------GCGTTGCCGCCGGC 839967

Qy       1027 aProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProPr 1047
              |||||||            |||   |||                  |||   :::|||
Db     839966 CCCGCCATTGCCACCGACCGCGCCGGCACCACCGCTTCCGCCGGCCCCGCCGGAGCCGAA 839907

Qy       1047 oCysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerPr 1067
                        :::            |||:::|||||||||   |||            ||
Db     839906 GAGCAAGGCGGCTCCGCCGGCCCCGCCAATCCCGCCGAGGCCGGTGGCGCCGGGCCCGCC 839847

Qy       1067 oHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGl 1087
              |      |||   |||         :::      ||||||      |||   |||
Db     839846 GCTGCCCCCGGTGCCGCCGGTGCCGAACAAGATCCCGCCGGCCCCGCCGGCCCCGCCGGC 839787

Qy       1087 y------LeuHisAspThrAlaArgProValLeuProArgProProThrIleSerAsnPr 1105
                     |||               |||      |||   ||||||      :::|||
Db     839786 CCCGCCGTTGACGCCGCCGATCGTTCCCTGCCCGCCGTTTCCGCCGTTGCCGAACAACCA 839727

Qy       1105 oProProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIl 1125
               ||||||         ||||||      |||   |||            :::||||||::
Db     839726 GCCGCCGGCTCCGCCGTCAGCCCCGGTCCCCGGAGTGCC----------GTTGGCGCCGT 839677

Qy       1125 eSerGln------------GlyMetSerValGlnLeuHisValProTyrSerGluHisAl 1141
              ::::               |||      ::::::      |||   :::   :::|||
Db     839676 TGCCGATCAGCGGACGCCCGGTGAGCGTCTGGAAGGGGTCGTTCACCATATTGATCACGT 839617

Qy       1141 a----------------------------------------------------------- 1141

Db     839616 CCTGCTGCAGGGTGTGCAGGGGCGAGGTGCTGGCGGGGGCCGTAAAGCCGTCCAGGCCCA 839557

Qy       1142 -LysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLysLe 1161
               :::   |||   |||            ||||||
Db     839556 ACAGCACCCCCCCGGTACCGCCGATGCCGGCCTTGCCCAGGGTCGCGCCGGTCCCGCCGC 839497
```

```
Qy       1161 uAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProPr 1181
               |||      :::   |||      :::               |||   |||||||||
Db     839496 TGCCGCCGTTGCCGCCGTTGCCGATCAACACGCCGGTGCCGCCGGCCCCGCCGGGCCCAC 839437

Qy       1181 oGluSerLeuGlyValProThrAla---GlnGluAlaSerValLeuArgGlyThrAlaLe 1200
                           ||||||   |||         |||:::|||            |||::
Db     839436 C-----------GTCCCCGCCGCCGGTGCTCGCGCCGCCGTTGCCGCCGTTGCCGCCGT 839389

Qy       1200 uGlySerValProGlyGlySerIleThrLys---------GlyIleProSerThrArgVa 1217
              :         |||||||||                        :::|||
Db     839388 TGCCGATCAGCCCGGGGGCCCCGCCGGCCCCGCCGGCGGCGCTGTCCCCGACGCTTTTAC 839329

Qy       1217 lProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVa 1237
                  :::      |||            |||:::      |||         |||
Db     839328 TAATGCCGCCGGCGCCGCCGGAGCCGCCGGAGCCGAAGAACATGCCGGCGTCGCC----- 839274

Qy       1237 lLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspAr 1257
                                             |||         |||:::
Db     839273 ---------------------------GCCGGCTCCGCCGGCCCCGCCGAATCCGCCGC 839242

Qy       1257 gGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLys----------- 1273
                           :::   |||   |||      |||      |||:::
Db     839241 TGCGACTAGCGCCGCCGGTCCCGCCGGCCCCGCCATCACCAGTGGCGAGTGTCCCGGCGT 839182

Qy       1274 ----LysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
                     |||               :::   |||            :::|||
Db     839181 TGCCGCCGGCCCCGCCGG-CCCCGCCAGTGAGGTTGCCCGACCCGCCGCTGCCGCGTTCC 839123

Qy       1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
                     |||   ::::::   ||||||               |||:::|||   :::
Db     839122 CTTCCGGCGCCAAACAGCCCGCCCGCCCCGCCGGCCCCGCCTGCCCCGCCGGTGGCAC-- 839065

Qy       1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
                           |||:::   |||
Db     839064 ----CGCCGTTCGAGAATCCGCCGACGCCGCCGACCCCGGCGGCGCCGAACAGCATGCCG 839009

Qy       1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
               |||         |||   |||      ||||||      :::   :::         ||
Db     839008 GCGTTTCCGCCGGCCCCGCCGG------CCCCACCGGCCCCGCCGGCGACGTCGGTGGAG 838955

Qy       1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
              |                     ||||||               :::
Db     838954 G---------------------CCCCGCCGGCCCCGCCGCTGCCGAACAGCCCGCCGGCC 838916

Qy       1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
               |||      :::      :::|||   |||   |||   ||||||   |||
Db     838915 CCGCCGTTACCGCCGGCCCCAGCGCCG-GGGGCGCCAGACCCGCCGGCGCCG-------- 838865

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
                                                ||||||
Db     838864 ----------------------------------CCGTTG-------------------- 838859

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
                                                   |||:::      |||
Db     838858 -------------------------------------CCGATCAACCACCCGCCGGGCCCGCC 838833

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
               |||         |||      ||||||||||||   |||:::
Db     838832 GTTGGCTCCGGTCCCGGGGGCGCCGTTGGCGCCGTTGCCGATC----------------- 838790

Qy       1452 rGlnGlyThrProLeu-------------------------------LysTyrAspTh 1461
                  |||   |||:::                                       |||||
Db     838789 -AGCGGGCGCCCGGTGGCCGCCACGAATTGCGCGTTGATCGGGGCCAGCAGCGGCGACAC 838731

Qy       1461 rGlyAlaSerThrThrGlySer-LysLysHisAspValArgSerLeuIleGlySerProG 1481
              |   |||::::::      :::    :::         ||||||               |
Db     838730 GGCGGCGGCCTCGGCGGCCGCATACGCGCCTCCGCCGGCGCTCAGGGCCTGCACGAACTG 838671

Qy       1481 lyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluA 1501
              |||||      ||||||      |||:::
Db     838670 GGCGTAAAACGCCGCCG------CCTGTG------------------------------ 838648

Qy       1501 rgAlaCysTyrGluGluSerLeuLysSerArgProGly-----------ThrAlaSerS 1517
                               |||         |||||||||            |||
Db     838647 --------------CGCTCGCCGCCTGGTAGGCCTGGCCTTGGGCACCGAACAACGCCGC 838602

Qy       1517 erSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProA 1537
                |||      :::      |||               :::|||         :::|||
Db     838601 AATCGCTGTCGACACCTCATCGGCACCCGCGGTCAACAATCCCGTCGTGTTGACGGCCG- 838543

Qy       1537 rgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuPro---- 1555
                                                             |||   |||
Db     838542 -----------------------------------------CCGCCGCATTAGCCGAGCT 838524

Qy       1556 -----------ArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluG 1572
                         |||      |||      :::      |||   |||         :::
Db     838523 AATCGCCGACCCGATCCCCGCCAAATCCGCTGCCGCCGCCACCACCAGCTCCGGCGAAAC 838464

Qy       1572 lySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSer------------- 1587
                |||      :::|||   |||||||||         ||||||:::
Db     838463 CATCACCCATGACATCGCATGCCTCCCAA------CCCCTCACGGCCTGGTGCCCGACTA 838410

Qy       1588 --------ThrProArgGluIleAla-----------------------------LysS 1595
                      ||||||      :::                                    :
Db     838409 CCAGCGACACACCCGTACGGGTAGGAATTAATTGCAGAAATCAACGATATCGCGATCCGG 838350

Qy       1595 erProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuL 1615
              ::||||||:::   :::|||                     :::|||            |
Db     838349 CACCGCATGCAGCGCTTCCACACGTTTCGCAG---------GCGCCAGCTGCTGGACGGC 838299

Qy       1615 euArgGlyValSerGlyValAspLeuTyrArgSerHisIlePro---------LeuAlaP 1632
              ||   |||      ||||||         |||:::   :::|||            |||
Db     838298 TCACGGGTCGCGGCGGCGTTGCCGACAAACGGGCGGCGGTGCCGGGCACAGGGCGCGCCG 838239

Qy       1632 heAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuP 1652
                   |||         |||   |||   |||               |||
Db     838238 CGAAACCCGCA------CCGCCCGGATAGCCGTCGTCACGACGACCCGCG---------- 838195

Qy       1652 roArgHisLeuAlaProAsnProThrTyrProHisLeuTyr-ProProTyrLeuIleArg 1671
                      |||      :::         |||   :::::: |||   ||||||   |||
Db     838194 --ACGGGCCTCTGACAAAGCGCTCGAGCGCCGGGGATCTTCGCCGAGATACCTGGCGCGG 838137

Qy       1672 GlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThr 1691
                    |||      ||||||                                    :::
Db     838136 CG-GCCCCCATTTGGCGGCCCGC-------------------------------TCC 838111

Qy       1692 SerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArg 1711
              :::      :::||||||                     :::               |||
Db     838110 GCAGCCGATCTGCATCACCGGCGAGTGGTTGACGCCCGGAAGCCATGCCGTGCCGCCCGG 838051

Qy       1712 GlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGly-Pro------Ar 1729
                       |||   |||      ||||||            |||||| |||
Db     838050 CTACGGCAACCCGTTGAGCCCGTTTTGGCCCAGCAGCAACCCGCCGGCGCCGCCGGTTCC 837991

Qy       1729 gGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGl 1749
               |||:::      |||:::   ||||||   ||||||   |||   ||||||   |||
Db     837990 GGGGCTGCCGGCGTTGGCGCCGGTGCCGCCGTTGCCGCCGTTG---CCGCCGGTGCCGAT 837934

Qy       1749 yThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSe 1769
               :::||||||                           ||||||   |||   |||
Db     837933 CAGCCCGGCGCTGCCG---------------------CCGACGCCACCGGCCCCG----- 837900

Qy       1769 rSerArgHisSerSerSerProLeuSerProGlyGlyProThr-His---------LeuT 1786
                  :::   |||         |||   |||      |||||| |||            |
Db     837899 -CCGAAGCCCAGCCAGCCGGCATTGCCGCCGGCCCCACCGACCCACCGGCGGTGCCGAA 837841

Qy       1786 hrLysProThr------------------ThrThrSerSerSerGluArgGluArgAspA 1800
              ||:::   |||                  ||||||::::::            |||   |
Db     837840 CCCGCCGAACCCGCCGGCCCCGCCGCTGGACAACAGCAGCCCGGCGTTGCCACCGGCCCC 837781

Qy       1800 rgAspArgGluArgAsp---------ArgAspArgGluArgGluLysSerIleLeuThrS 1817
              ||         |||            |||   |||   |||   :::         |||:
Db     837780 GCCGTCCCCACCGAGGAAACCAAATCCGCCGGCGCCGCCGGCGCCGCCGCTGCCGAACAG 837721

Qy       1817 erThrThrThrValGlu------HisAlaProIleTrpArgProGlyThrGluGln-Ser 1834
              ::                           |||         ||||||      ::: :::
```

```
Db     837720 CAGCCCTGCGCTACCGCCGGCCCCGCCGTCCCCACCGGTGTCCAGGTTTTCGCCGTCACC 837661

Qy       1835 SerGlySerSerGlySerSer------GlyGlyGlyGlyGlySerSerSerArgProAla 1852
              :::||||||:::|||::::::      |||         |||      :::   ||||||
Db     837660 GCCGGCTCCGCCGGCGCCGCCCGGCCCGGCGAGCATTCCGGCATTCCCGCCGGCCCCGCC 837601

Qy       1853 SerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGln 1872
              :::   :::   |||            |||      |||            |||
Db     837600 AACCCCGCCGTCGCCATTAGTGCTGAGCCCGCCGGTCCCGCCGTGACCTCCGCCGGCACC 837541

Qy       1873 GlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPro 1892
                 :::|||            :::   |||      |||                  |||
Db     837540 GATCAGCCCGGC---CAGCAGTCCGCCGGTCCCGCCGGC------------------CCC 837502

Qy       1893 SerLysProThrValLeuArgSerThrSerThrSerSer-ProValArgProAlaAlaTh 1912
              :::   |||   |||            :::|||:::    ||||||   ||||||
Db     837501 ACCGGCCCCGCCGTC------------GCCACCAACAAGGCCGGTTCCACCGGCCCCGCC 837454

Qy       1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
                  |||||||||   |||   |||                           ||||||
Db     837453 GGCCCCACCGGCGCCCCACAGCCCA--------------------------CCGACCCC 837421

Qy       1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
                     |||   :::   |||               |||      |||      |||
Db     837420 GCCTCCCCCGCCGGTGCCGCCGAGT---------ACGGTGCTGGCTCCGCCGATCCCGCC 837370

Qy       1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               |||            :::      ::::::|||||||||               |||||
Db     837369 GGCCCCGCCGTCGCCGAGC------AGCCAGCCGCCGGCCCCGCCGGCC---CCACCGGC 837319

Qy       1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
              |      |||               |||   |||      |||||||||      |||::
Db     837318 CCCGCCGCCACCCGCCGAGCTGCCTCCGGCCCCGCCGGCCCCGCCGGTACCC---CACAG 837262

Qy       1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
              :      |||      ||||||         ||||||      |||:::|||||||||||
Db     837261 CCCG---GCAGCCCCGCCCGCCCCGCCGGGCGCGCCCGAGCCCGCCGCGCCGGACCCGCC 837205

Qy       2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
              ||||||||||:::   :::|||   |||
Db     837204 GGCCCCGCCGTCGCCGAGCAGCCACCCA-------------------------------- 837177

Qy       2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052

Db     837177 ------------------------------------------------------------ 837177

Qy       2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
                     |||   :::||||||:::         |||            |||:::|||
Db     837176 -------CCGGGGGCCCCGGTGGCC------CCGCTGCCGACCGCCCCGGGGGTGCCG-- 837132

Qy       2072 sHisLeu-GluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProG 2092
                  |||                      :::   |||      |||      :::|||
Db     837131 ----TTGCCGATCAGCGGTCGCCCGGTCAGTGTCTGGGTCGGCGCGTTGATCGCGGCCAG 837076

Qy       2092 lyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluS 2112
                ||||||   :::   |||   ||||||                     |||      |
Db     837075 CGCCTGTT---GTTTCAGGGTGTGCAGCGGAGAG-GCGCTGGCCGGGGTGTTGAAGCCGT 837020

Qy       2112 erGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgV 2132
              ||   ||||||||||||||||||:::      ||||||
Db     837019 CGGCGCCCAGCAGCAGCCCACTGATCCCACCCGCGCCG---------------------- 836982

Qy       2132 alValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisP 2152
                      ::::::      :::      |||                           |
Db     836981 --------GTATCACCCGCGGTCGGTCCGGTT---------------------------CCGC 836954

Qy       2152 roGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysP 2172
              ||               |||||||||            |||   |||            |
Db     836953 CGATGCCGGCGTTGCCGCCGTTGCCGATCAGCACCCCGTCGCCGCCGGCCCCGCCGGTAC 836894

Qy       2172 roValLeuAspLeuArgArgProPro---------------------------SerAspL 2183
              ||   :::            ||||||                           |||:::
Db     836893 CGCCGGTCACGGCGCCTTGCCCACCGGCGCCGCCGGCCCCGCCGTTACCGATCAGCAACC 836834

Qy       2183 euTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyG 2203
                   |||||||||               |||:::      ||||||:::
Db     836833 CGGCCTTGCCGCCGGCCCCGCCGACACCCCCGTCACCGGTTTCGCCGAACCCGCCGGCCC 836774

Qy       2203 lyLysArgSerProGluProAsnLysThrSerVal---------LeuGlyGlyGlyGluA 2220
                         ||||||||||||:::      |||               |||
Db     836773 CGCCGACCCCACCGGAGCCCAACCAGCC-GGCGTCACCGCCGGCACCGCCGGCTCCCCCG 836715

Qy       2220 spGlyIle---GluProValSerProProGluGlyMetThrGluProGlyHisSerArgS 2239
                ||||||   ||||||      ||||||                     |||:::
Db     836714 TCGGTATCGCCGAACCCGCCGGTCCCCCGGATCCGGCGTTGCCGAACAGCACACCGGTG 836655

Qy       2239 erAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSer---------- 2255
                |||      |||                        |||   |||:::
Db     836654 TTGCCGCCGGCCCC------------------------ACCGGCCCCACCGTCACCACCT 836619

Qy       2256 --------ArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheS 2273
                            ||||||               |||   |||   |||         :
Db     836618 AGGCCGGCTCCGCCGGATCCACCGGTGCCGCCGGCACCGAACAGCAGCCCGGCGTTGCCG 836559

Qy       2273 erLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysL 2293
              ::      |||:::                                    :::   :::|
Db     836558 CCAGCCCCACCGACCCCACC------------------------------GTTGTTGAGA 836529

Qy       2293 ysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly------- 2310
              ||      :::               |||               |||||||||
Db     836528 AAGCCACGTCCGCCGGTCCCGCCGGCCCCGCCCGGCCCGGCCAGCAGCCCGGCCAGCAGC 836469

Qy       2311 --ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerG 2330
                |||         :::   |||      ||||||   |||   :::      |||   :
Db     836468 CCACCGGCTCCGCCAGTTCCCCCGACCCCACCGGTGACGGCGC-CTAGCCCGCCGGCGCC 836410

Qy       2330 lnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaL 2350
              ::      ::::::|||
Db     836409 GCCGTGCCCGCCAGCAC------------------------------------------- 836393

Qy       2350 euMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSer-AlaAsnAlaPheAsn 2369
                                                 |||   ::: |||         |||
Db     836392 -----------------------------------CCAACAGCCCGCCGGCCCCACCAAC 836368

Qy       2370 ProLeuAsnAlaSerAlaSerLeuPro------------------------------Ala 2379
              |||         :::   :::   |||                              |||
Db     836367 CCCGCCGGCCCCACCGGCGCCGGCCCCGAGCCCGCCGACCCCGCCGACCCCACCGGTGCC 836308

Qy       2380 AlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeu--------------- 2394
                    |||   |||   :::   |||               :::
Db     836307 CAACAACCAGCCACCTGATCCACCGGCACCCCCGGCCTCGCCATCACCTACCGTGGAACT 836248

Qy       2395 ------------ThrSerProGlyGly------------GlyGlyLysAlaLysValSer 2406
                          |||||||||   |||            ||||||   |||      :::
Db     836247 CCCGCCGGCCCCACCAGCCCCGCCGGTGCCGAACAACCCGGCGGCCCCGCCGGCCCCGCC 836188

Qy       2407 -GlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspAr 2426
               |||||||||            |||      |||||||||   |||   |||   :::
Db     836187 GGGCAGGCCCGAGCCAGCGGCGGACCCGCCGGCCCCACCATTACCGAGC---AACCA 836131

Qy       2426 gProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuTh 2446
               ||||||                        |||                  ||||||
Db     836130 CCCGCCCGGCGCACCATTGGCCCCCGTGCCCGGG---------------GCCCCGTTGGC 836086

Qy       2446 rAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPhe---ProTyrAs 2465
                                 ||||||   ||||||                        ::
Db     836085 GCCGTTGCCGATCAACGGGCGCCCCGTGAGCGCGACACTTTGCGCATTGACGGCGGCCAG 836026

Qy       2465 nProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuPr 2485
              :               |||   ||||||:::   ::::::      |||   ||||||
Db     836025 CAGGTCCTGCTGCACGCTCTGCGCCGGGTTGGCCAGCACCGCCGCGCCGGC-GGCCTCCG 835967

Qy       2485 oAlaGlySerGlyPro 2490
               ||||||   ||||||
Db     835966 CGCTGGCGTAGGCCCC 835951
```

```
RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:               1.54e-19       Length:          4403765
Score:                   685.00         Matches:         530
Percent Similarity:      32.37%         Conservative:    186
Best Local Similarity:   23.96%         Mismatches:      852
Query Match:             5.18%          Indels:          652
DB:                      3              Gaps:            95

US-09-522-753-5 (1-2517) x US-09-103-840A-2 (1-4403765)

Qy        546 LeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValAla 565
              ::::::   |||               |||:::      ||||||   |||   :::
Db     843605 GTTCGTCGGAAGGTGATGGGGCTGGGCGGTGACGGCTTCGACGAATGCGAATGCCTCTGC 843546

Qy        566 SerLysGlyArgLysThrAlaAsnSer------------------GlnGlyArgArg--- 578
              |||:::   |||         ::::::                  :::   ||||||
Db     843545 TCGCGGTGACGGAATCTCGAAGATGCGTCGATTCGTTGCCAGCCGGAGGAACGACGCCCA 843486

Qy        579 ------------------LysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSer 592
                                :::|||   ::::::|||      :::   |||      :::
Db     843485 TACCAGGTTCGGCACTGTGAAGGGGTCGTCAGCCGCGAGCAGTCGATCGAACCAGGGGCG 843426

Qy        593 Glu---------------------GluAlaIleThrProGlnGln--------------- 600
              :::                        |||:::      ||||||
Db     843425 GACGGTTCGGTGATTCGGGTGGTCACCGCGGTGTGCAGCCAGCAGCACGTTGGCGTCGAG 843366

Qy        601 ---------------SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrp 615
                             ||||||   :::                     |||||||||
Db     843365 AAGGAACATTGCCGTTCTGCGCCTATTGGGGGTTACTTCCGGGAGTTCAGTTCCAGGCCC 843306

Qy        616 ThrGluGluGluMetGluThrAla------------------------------------ 623
              :::            :::   |||
Db     843305 TCGTCGAGCACTTCGGACAACGCTCTATTCGAGGTCAGGTCGATGCCTCGCCGCGGACCG 843246

Qy        624 ------------LysLysGlyLeuLeuGluHisGlyArg---------AsnTrpSerAla 636
                          |||   |||   :::      ||||||            |||:::|||
Db     843245 GTGCCGCCGTCGAAAACCGGGACGGTC------GGGCGGGCGCCGCCGACGTGGGCGGCG 843192

Qy        637 Ile--------------------------------------------------------- 637

Db     843191 GCGAACTCCCGCCGAAGGGCGTCCTCGATCACAGCGCCCAGCGATTGACCACGCTCGCGG 843132

Qy        638 AlaArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLys 657
              ||||||   :::   :::   ::::::|||
Db     843131 GCCCGGCGTTTGGCGGCAGCGAGTATTTCA------------------------------ 843102

Qy        658 LysArgGlnAsnLeuAspGluIleLeu------------------------GlnGlnHis 669
                          :::|||   ::::::                        |||||||||
Db     843101 ------TCGGAGATTGACACCGTGGTGCGCATGATGCACACGGTAACGCATCAACAGCAT 843048

Qy        670 LysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAla 689
                 |||:::   |||:::   :::         :::::::::   ::::::   |||
Db     843047 CATCTGCGGTC-GAGCGGCCGAAGCGATGCTGTCAACGCCGCGCGCGGTCGCAGGCCCGC 842989

Qy        690 AlaSerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGly 709
                       :::|||      |||   :::   |||   :::         |||:::|||
Db     842988 CGGCTACATCAAGCCGTTCAGCCCGTTCTCACCGAGCAGCAGCCCGCCGGCGCCGCCGGC 842929

Qy        710 ValSerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsn 729
                 :::|||
Db     842928 GCCGCTGGG--------------------------------------------------- 842920

Qy        730 GluValProArgGlyGluCysSer------------GlyProAlaThrValAsnAsnSer 745
                                |||:::            |||         |||
Db     842919 ------------GCCGGGTGCACCCCCGGATTTCCCGGCGTTACCGCTGTTACCGCCGTT 842872

Qy        746 SerAspThrGluSerIleProSerProHisThrGluAlaAla------------------ 759
              :::         ::::::         |||      ||||||
Db     842871 GCCGCCGTTGCCGCCGTTGCCGATCAGCACGGCGCTGCCGCCGACGCCGCCCTTGCCGCC 842812

Qy        760 ------LysAspThrGlyGlnAsnGly-------ProLysProProAlaThrLeuGlyAl 775
                    :::   |||||||||   |||      |||   ||||||:::
Db     842811 GATGTGGAAAGCACCGGACAAACCGGTGGCGCCGCCGGCGCCGCCGTCACCTCCGAGACC 842752

Qy        775 aAspGlyPro------ProProGlyProProThrProPro------ArgArgThrSerAr 791
               :::   |||      ||||||   ||||||:::||||||               :::
Db     842751 GAACGCTCCGGACTTGCCGCCGGCCCCGCCGTCCCCGCCGGTGGCGGCGCCGGTGGCGGA 842692

Qy        791 gAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProProAl 811
               ||||||      |||   |||      |||   |||      |||      ||||||||
Db     842691 GGCGCCGGCCCCGCCGGCCCCG---CCGGAGCCGACGAGCAAACCCGCGTCGCCGCCGGC 842635

Qy        811 aProProSerProSerAlaPro------------ProProValValProLysGluGluLy 827
              |||||||:::|||      |||            ||||||:::   |||         ::
Db     842634 CCCGCCGGCCCCGCCGTTGCCAGAGTTGCTGAACCCGCCGCTCCCGCCGGCGCCGCCGGA 842575

Qy        827 sGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLysProProAl 847
              :   :::            |||||||||:::                        |||
Db     842574 GCCGAAGAGCAAGCCTCCGGCGCCGCCGATG------------------------CCGCC 842539

Qy        847 aAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysTh 867
                           ||||||      |||   |||
Db     842538 GACCACGCCGGCGGCGGTCAG-GCTGGAACCGCC-------------------------- 842506

Qy        867 rGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAl 887
                     |||      |||||||||   |||   :::   |||   |||         ||
Db     842505 -------GCTGCCGCCGGCCCCGCCGGAGGCGCCGAGGGAGAGCAAGCCGGCGTCGCCGC 842453

Qy        887 aGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLy 907
              |      |||      :::            ||||||:::      |||      |||
Db     842452 CGGCCCCGCCGGCCCCTCCGGAAACTCCGGCGGCGGTGCTATGGCTGCCGGCGCCGCCGG 842393

Qy        907 sSerSerGlyAlaProGln------------------------------AspSerAspSe 917
               ::::::||||||   |||                              :::      ::
Db     842392 TGCCGCCGGCGCCGAACAGCCCGCCGTTGCCGCCCGCCCCGCCGGCGCCGAACAGCCCGC 842333

Qy        917 rSerAlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLe 937
              :      ||||||   ||||||
Db     842332 CGTTCCCACCTGCCCCGCCGATCCCGCC-------------------------------- 842305

Qy        937 uLeuSerProArgProSerLeuLeuThrProThrGlyAsp-ProArgAlaAsnAlaSerP 957
                                       :::|||:::|||    |||               |
Db     842304 -------------------------AGCCCCAGCGGGGCTGCCAAACCCACCGGGCCCGC 842270

Qy        957 roGlnLysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleG 977
              ||      ||||||                           |||      ||||||:::
Db     842269 CGGCTCCCCCGTTG---------------------------GCGAACAGCCCGCCGTTA- 842238

Qy        977 lnValThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProP 997
                                  ||||||               ||||||   ||||||||||
Db   842237 --------------------CCGCCGGCCCCGCCATTGCCACCGACCGCGCCGGCACCAC 842198

Qy      997 roAlaProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlyS 1017
            |||||||||||                              ||||||      |||   |
Db   842197 CGGCCCCGCCG------------------------------GCCCCGGCGGCGCCCAAGA 842168

Qy     1017 erSerProArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaA 1037
            ||:::|||         :::      ||||||||||||            |||   |||
Db   842167 GCGCACCG---------GCGTTGCCGCCGGCCCCGCCATTGCCACCGACCGCGCCGGCAC 842117

Qy     1037 laGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProV 1057
                             |||   :::|||            :::            |||:
Db   842116 CACCGCTTCCGCCGGCCCCGCCGGAGCCGAAGAGCAAGGCGGCTCCGCCGGCCCCGCCAA 842057

Qy     1057 alProProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerT 1077
            ::|||||||||   |||            |||      |||   |||         :::
Db   842056 TCCCGCCGAGGCCGGTGGCGCCGGGCCCGCCGCTGCCCCCGGTGCCGCCGGTGCCGAACA 841997

Qy     1077 yrAlaProProGlyHisProLeuProLeuGly------LeuHisAspThrAlaArgProV 1095
                 ||||||      |||   |||            |||               |||
Db   841996 AGATCCCGCCGGCCCCGCCGGCCCCGCCGGCCCCGCCGTTGACGCCGCCGATCGTTCCCT 841937

Qy     1095 alLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisP 1115
                 |||   ||||||      :::|||   ||||||         ||||||      |
Db   841936 GCCCGCCGTTTCCGCCGTTGCCGAACAACCAGCCGCCGGCTCCGCCGTCAGCCCCGGTCC 841877

Qy     1115 roSerValLeuGluArgGlnIle------------------------------------- 1122
            ||   |||            :::
Db   841876 CCGGAGTGCCGTTGGCGCCGTTGCCGATCAGCGGACGCCCGGTGAGCGTCTGGAAGGGGT 841817

Qy     1123 -----------------------------------------GlyAlaIleSerGlnGlyM 1129
                                                     |||   :::      |||
Db   841816 CGTTCACCATATTGATCACGTCCTGCTGCAGGGTGTGCAGGGGCGAGGTGCTGGCGGGG- 841758

Qy     1129 etSerValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValT 1149
                 |||:::         |||   |||               |||:::      :::
Db   841757 --GGCGTAAAGCCGTCCAGGCCCAACAGCACCCCCCCGGTACCGCCGATGCCGGCCTTGC 841700

Qy     1149 hrMetGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysG 1169
                       |||:::|||      |||            |||         :::
Db   841699 CCAGGGTCGCGCCGGTCCCGCCGCYGCCG------------CCGTTGCCGCCGTTGCCGA 841652

Qy     1169 lnGluGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrA 1189
                       |||      |||   |||||||||               ||||||   |
Db   841651 TCAACACGCCGAGTGCCGCCGGCCCCGCCGGGCCCACC------------GTCCCCGCCG 841604

Qy     1189 la---GlnGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerI 1208
            ||         |||:::|||            |||:::         |||||||||
Db   841603 CCGGTGCTCGCGCCGCCGTTGCCGCCGTTGCCGCCGTTGCCGATCAACCCGGGGGCCCCG 841544

Qy     1208 leThrLys---------GlyIleProSerThrArgValProSerAspSerAlaIleThrT 1225
                                :::|||               :::      |||
Db   841543 CCGGCCCCGCCGGCGGCGCCGTCCCCGACGCTTCTACTAATGCCGCCGGCGCCGCCGGAG 841484

Qy     1225 yrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrA 1245
                 |||:::      |||         |||
Db   841483 CCGCCGGAGCCGAAGAACATGCCGGCGTCGCC---------------------------- 841452

Qy     1245 rgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProL 1265
                    |||         |||:::                        :::   |||
Db   841451 -----GCCGGCTCCGCCTGCCCCGCCGAATCCGCCGCTGCGACTAGCGCCGCCGGTCCCG 841397

Qy     1265 ysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetS 1285
              |||      |||      |||::::::      |||   :::
Db   841396 CCGGCCCCGCCATCACCAGTGGCGAGCGTCCCGGCGTTGCCGCC---------------- 841353

Qy     1285 erValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrA 1305
                                                  :::||||||      |||   |
Db   841352 --------------------------------GGCCCCGCCGGCCCCGCCAGTGAGGTTG 841325

Qy     1305 laAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerA 1325
            ||   |||                                             ::::::|
Db   841324 CCCGACCC---------------------------------------------GCCGCTG 841310
```

```
Qy     1325 laSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerProHisHisL 1345
            ||:::                     |||   ||||||   |||      |||
Db   841309 CCGCCTTCCCCTCCGG-CGCCAAACAGCCCGCCCGCCCCGCCGG------CCCCGCCTGC 841257

Qy     1345 euLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSerTyrValG 1365
              :::      |||                                 ||||||   :::
Db   841256 CCCGCCGGTGGCACCGC------------------------------CGTTCGAGAATCC 841227

Qy     1365 luAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProP 1385
                 :::            ||||||         :::|||               ||||
Db   841226 GCCGACGCCGCCGACCCCGGCGGCGCCGAAC-AGCATGCTGGCGTTTCCGCCGGCCCCGC 841168

Qy     1385 ro---ProProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLeu---- 1402
            ||   ||||||   |||               ||||||               |||
Db   841167 CCGCCCCACCGGCCCCGCCGGCGACGTCGGTGGAGGCCCCGCCGGCCCCGCCGCTGCCGA 841108

Qy     1403 --------------GlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrV 1418
                             ||||||         ||||||
Db   841107 ACAGCCCGCCGGCCCCGCCGTTACCGCCGGCCCCAGCG---------------------- 841070

Qy     1418 alLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProG 1438
                       |||                  |||         :::   |||   |||
Db   841069 --------CCGGGGGCGCCAGACCCGCCGGCGCCGCCGTTGCCGATCAACCACCCGCCGG 841018

Qy     1438 luLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeu- 1457
                 ||||||||||||   |||                        |||   ||||||
Db   841017 GCCCGCCGTTGGCTCCGGTCCCG------------------------GGGGCGCCGTTGG 840982

Qy     1457 ------------------------------------------------------------ 1457

Db   840981 CGCCGTTGCCGATCAGCGGGCGCCCGGTGGCCGCCACGAATTGCGCGTTGATCGGGGCCA 840922

Qy     1458 --LysTyrAspThrGlyAlaSerThrThrGlySer-LysLysHisAspValArgSerLeu 1476
                    ||||||   |||::::::      :::    :::         ||||||
Db   840921 GCAGCGGCGACACGGCGGCGGCCTCGGCGGCCGCATACGCGCCTCCGCCGGCGCTCAGGG 840862

Qy     1477 IleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAla 1496
                        ||||||      ||||||      |||:::
Db   840861 CCTGCACGAACTGGGCGTAAAACGCCGCCG------CCTGTG------------------ 840826

Qy     1497 ArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGly--------- 1513
                                          |||         |||||||||
Db   840825 ---------------------------CGCTCGCCGCCTGGTAGGCCTGGCCTTGGCAC 840793

Qy     1514 ---ThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGlu 1532
               |||         |||      :::      |||               :::|||
Db   840792 CGAACAACGCCGCAATCGCTGTCGACACCTCATCGGCACCCGCGGTCAACAATCCCGTCG 840733

Qy     1533 LeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGly 1552
                  :::|||
Db   840732 TGTTGACGGCCG------------------------------------------CCGCCG 840715

Qy     1553 HisLeuPro---------------ArgGlySerProValThrMetArgGluProThrPro 1567
            |||   |||               |||      |||      :::      |||   |||
Db   840714 CATTAGCCGAGCTAATCGCCGACCCGATCCCCGCCAAATCCGCTGCCGCCGCCACCACCA 840655

Qy     1568 ArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSer 1587
                     :::   |||      :::|||   |||||||||         ||||||:::
Db   840654 GCTCCGGCGAAACCATCACCCATGACATCGCATGCCTCCCAA------CCCCTCACGGCC 840601

Qy     1588 ---------------------ThrProArgGluIleAla--------------------- 1593
                                 ||||||      :::
Db   840600 TGGTGCCCGACTACCAGCGACACACCCGTACGGGTAGGAATTAATTGCAGAAATCAACGA 840541

Qy     1594 ---------LysSerProHisSerThrValProGluHisHisProHisProIleSerPro 1610
                        :::||||||:::   :::|||                     :::|||
Db   840540 TATCGCGATCCGGCACCGCATGCAGCGCTTCCACACGTTTCGCAG---------GCGCCA 840490

Qy     1611 TyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIlePro--- 1629
                        |||   |||      ||||||         |||:::   :::|||
Db   840489 GCTGCTGGACGGCTCACGGGTCGCGGCGGCGTTGCCGACAAACGGGCGGCGGTGCCGGGC 840430
```

```
Qy       1630 ------LeuAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAla 1647
                       |||      |||         |||   |||   |||
Db     840429 ACAGGGCGCGCCGCGAAACCCGCA------CCGCCCGGATAGCCGTCGTCACGACGACCC 840376

Qy       1648 AlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyr-ProPr 1667
              |||                  |||      :::         |||   :::::: |||
Db     840375 GCG------------ACGGGCCTCTGACAAAGCGCTCGAGCGCCGGGGATCTTCGCCGAG 840328

Qy       1667 oTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAs 1687
               ||||||   |||      |||      ||||||
Db     840327 ATACCTGGCGCGGCG-GCCCCCATTTGGCGGCCCGC------------------------ 840292

Qy       1687 nAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAl 1707
                        ::::::      :::||||||                     :::
Db     840291 ----------TCCGCAGCCGATCTGCATCACCGGCGAGTGGTTGACGCCCGGAAGCCATG 840242

Qy       1707 aAspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGl 1727
                        |||         |||   |||      ||||||            |||||
Db     840241 CCGTGCCGCCCGGCTACGGCAACCCGTTGAGCCCGTTTGGCCCAGCAGCAACCCGCCGG 840182

Qy       1727 y-Pro------ArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValP 1745
              | |||         |||:::      |||:::   ||||||   ||||||   |||   |
Db     840181 CGCCGCCGGTTCCGGGGCTGCCGGCGTTGGCGCCGGTGCCGCCGTTGCCGCCGTTG---C 840125

Qy       1745 roProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaP 1765
              |||||   |||   :::||||||                           ||||||   |
Db     840124 CGCCGGTGCCGATCAGCCCGGCGCTGCCG---------------------CCGACGCCAC 840086

Qy       1765 roGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGlyProThr-His 1784
              ||   |||         :::   |||         |||   |||      |||||| |||
Db     840085 CGGCCCCG------CCGAAGCCAGCCAGCCGGCATTGCCGCCGGCCCCACCGACCCCAC 840032

Qy       1785 ---------LeuThrLysProThr------------------ThrThrSerSerSerGlu 1795
                          |||:::   |||                  ||||||::::::
Db     840031 CGGCGGTGCCGAACCCGCCGAACCCGCCGGCCCCGCCGCTGGACAACAGCAGCCCGGCGT 839972

Qy       1796 ArgGluArgAspArgAspArgGluArgAsp---------ArgAspArgGluArgGluLys 1812
                    |||   |||         |||            |||   |||   |||   :::
Db     839971 TGCCACCGGCCCCGCCGTCCCCACCGAGGAAACCAAATCCGCCGGCGCCGCCGGCGCCGC 839912

Qy       1813 SerIleLeuThrSerThrThrThrValGlu------HisAlaProIleTrpArgProGly 1830
                       |||:::                           |||         ||||||
Db     839911 CGCTGCCGAACAGCAGCCCTGCGCTACCGCCGGCCCCGCCGTCCCCACCGGTGTCCAGGT 839852

Qy       1831 ThrGluGln-SerSerGlySerSerGlySerSer------GlyGlyGlyGlyGlySerSe 1848
                    ::: ::::::||||||:::|||::::::      |||         |||
Db     839851 TTTCGCCGTCACCGCCGGCTCCGCCGGCGCCGCCCGGCCCGGCGAGCATTCCGGCATTCC 839792

Qy       1848 rSerArgProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGl 1868
               :::   ||||||:::   :::   |||            |||      |||
Db     839791 CGCCGGCCCCGCCAACCCCGCCGTCGCCATTAGTGCTGAGCCCGCCGGTCCCGCCGTGAC 839732

Qy       1868 nAspAlaLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleTh 1888
                  |||         :::|||            :::   |||      |||
Db     839731 CTCCGCCGGCACCGATCAGCCCGGC---CAGCAGTCCGCCGGTCCCGCCGGC-------- 839683

Qy       1888 rAlaValGluProSerLysProThrValLeuArgSerThrSerThrSerSer-ProValA 1908
                        |||:::   |||   |||            :::|||:::    ||||||
Db     839682 ----------CCCACCGGCCCCGCCGTC------------GCCACCAACAAGGCCGGTTC 839645

Qy       1908 rgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyV 1928
                ||||||         |||||||||   |||   |||
Db     839644 CACCGGCCCCGCCGGCCCCACCGGCGCCCCACAGTCCA---------------------- 839607

Qy       1928 alTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgP 1948
                   ||||||         |||   :::   |||               |||      |
Db     839606 -----CCGACCCCGCCTCCCCCGCCGGTGCCGCCGAGT---------ACGGTGCTGGCTC 839561

Qy       1948 roGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerG 1968
              ||      |||   |||            :::      ::::::|||||||||
Db     839560 CGCCGATCCCGCCGGCCCCGCCGTCGCCGAGC------AGCCAGCCGCCGGCCCCGCCGG 839507

Qy       1968 lyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProProV 1988
                      ||||||      |||               |||   |||      |||||||
Db     839506 CC---CCACCGGCCCCGCCGCCACCCGCCGAGCTGCCTCCGGCCCCGCCGGCCCCGCCGG 839450

Qy       1988 alSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaS 2008
              ||      |||:::      |||      ||||||         ||||||      |||:
Db     839449 TACCC---CACAGCCCG---GCAGCCCCGCCCGCCCCGCCGGGCGCGCCCGAGCCCGCCG 839396

Qy       2008 erProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnS 2028
              ::|||||||||||||||||||||:::   :::|||   |||
Db     839395 CGCCGGACCCGCCGGCCCCGCCGTCGCCGAGCAGCCACCCA------------------- 839355

Qy       2028 erLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerT 2048

Db     839355 ------------------------------------------------------------ 839355

Qy       2048 yrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAspL 2068
                                  |||   :::||||||:::         |||
Db     839354 --------------------CCGGGGGCCCCGGTGGCC------CCGCTGCCGACCGCCC 839321

Qy       2068 ysGlyLeuProLysHisLeu-GluGluLeuAspLysSerHisLeuGluGlyGluLeuArg 2087
                |||:::|||      |||                      :::   |||      |||
Db     839320 CGGGGGTGCCG------TTGCCGATCAGCGGTCGCCCGGTCAGTGTCTGGGTCGGCGCGT 839267

Qy       2088 ProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArg 2107
                    :::|||   ||||||   :::   |||   ||||||
Db     839266 TGATCGCGGCCAGCGCCTGTT---GTTTCAGGGTGTGCAGCGGAGAG-GCGCTGGCCGGG 839211

Qy       2108 ProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLys 2127
                 |||      |||   ||||||||||||||||||:::      ||||||
Db     839210 GCGTTGAAGCCGTCGGCGCCCAGCAGCAGCCCACTGATCCCACCCGCGCCG--------- 839160

Qy       2128 GlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyr 2147
                                   ::::::      :::      |||
Db     839159 ---------------------GTATCACCCGCGGTCGGTCCGGTT--------------- 839136

Qy       2148 ThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePro 2167
                          |||               |||||||||            |||   |||
Db     839135 ---------CCGCCGATGCCGGCGTTGCCGCCGTTGCCGATCAGCACCCCGTCGCCGCCG 839085

Qy       2168 GlyAlaSerCysProValLeuAspLeuArgArgProPro--------------------- 2180
                          |||   :::            ||||||
Db     839084 GCCCCGCCGGTACCGCCGGTCACGGCGCCTTGCCCACCGGCGCCGCCGGCCCCGCCGTTA 839025

Qy       2181 ------SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerPro 2198
                    |||:::      |||||||||               |||:::      ||||||
Db     839024 CCGATCAGCAACCCGGCCTTGCCGCCGGCCCCGCCGACACCCCCGTCACCGGTTTCGCCG 838965

Qy       2199 HisSerGluGlyGlyLysArgSerProGluProAsnLysThrSerVal---------Leu 2215
              :::                     ||||||||||||:::      |||
Db     838964 AACCCGCCGGCCCCGCCGACCCCACCGGAGCCCAACCAGCC-GGCGTCACCGCCGGCACC 838906

Qy       2216 GlyGlyGlyGluAspGlyIle---GluProValSerProProGluGlyMetThrGluPro 2234
                 |||         ||||||   ||||||      ||||||
Db     838905 GCCGGCTCCCCCGTCGGTATCGCCGAACCCGCCGGTCCCCCCGGATCCGGCGTTGCCGAA 838846

Qy       2235 GlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluPro 2254
                 |||:::      |||      |||                        |||   |||
Db     838845 CAGCACACCGGCGTTGCCGCCGGCCCC------------------------ACCGGCCCC 838810

Qy       2255 Ser------------------ArgMetGlySerLysSerProGlyAsnThrSerGlnPro 2268
              :::                        ||||||               |||   |||
Db     838809 ACCGTCACCACCTAGGCCGGCTCCGCCGGATCCACCGGTGCCGCCGGCACCGAACAGCAG 838750

Qy       2269 ProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGln 2288
              |||         :::      |||:::
Db     838749 CCCGGCGTTGCCGCCAGCCCCACCGACCCCACC--------------------------- 838717

Qy       2289 GluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGln 2308
                 :::   :::|||      :::               |||               |||
Db     838716 ---GTCGTTGAGAAAGCCACGTCCGCCGGTCCCGCCGGCCCCGCCCGGCCCGGCCATCAG 838660

Qy       2309 ProGly---------ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeu--- 2324
              ||||||         |||         :::   |||      ||||||   |||
Db     838659 CCCGGCCAGCAGCCCACCGGCTCCGCCAGCTCCCCGACCCCACCGGTGACGGCGCCGAG 838600

Qy       2325 ---------MetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341
                                   :::         |||            ||||||
Db     838599 CCCGCCGGCGCCGCCGTGCCCGCCAGCACCCAACAGCCCGCCGGCCCCACCAAC------ 838546

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerPro 2361

Db     838546 ------------------------------------------------------------ 838546

Qy       2362 ProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuPro------Ala 2379
              |||                     |||      |||      :::   |||      |||
Db     838545 CCCGCCGGCCCCACCGGCGCCGGCCCCGAGCCCGCCGACCCCGCCGACCCCACCGGTGCC 838486

Qy       2380 AlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeu--------------- 2394
                    |||   |||         |||               :::
Db     838485 CAACAACCAGCCACCTGACCCACCGGCACCCCCGGCCCCGCCATCACCTACCGTGGAACT 838426

Qy       2395 ------------ThrSerProGlyGly------------GlyGlyLysAlaLysValSer 2406
                          |||||||||   |||            ||||||   |||      :::
Db     838425 CCCGCCGGCCCCACCAGCCCCGCCGGTGCCGAACAACCCGGCGGCCCCGCCGGCCCCGCC 838366

Qy       2407 -GlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspAr 2426
               |||||||||            |||      |||||||||   |||   |||   :::
Db     838365 GGGCAGGCCCGAGCCAGCGGCGGCGGACCCGCCGGCCCCACCATTACCGAGC---AACCA 838309

Qy       2426 gProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuTh 2446
               ||||||                        |||                  ||||||
Db     838308 CCCGCCCGGCGCACCATTGGCCCCCGTGCCCGGG---------------GCCCCGTTGGC 838264

Qy       2446 rAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPhe---ProTyrAs 2465
                                 ||||||   ||||||                        ::
Db     838263 GCCGTTGCCGATCAACGGGCGCCCCGTGAGCGCGACACTTTGCGCATTGACGGCGGCCAG 838204

Qy       2465 nProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuPr 2485
              :               |||   ||||||:::   ::::::      |||   ||||||
Db     838203 CAGGTCCTGCTGCACGCTCTGCGCCGGGTTGGCCAGCACCGCCGCGCCGGC-GGCCTCCG 838145

Qy       2485 oAlaGlySerGlyPro 2490
               ||||||   ||||||
Db     838144 CGCTGGCGTAGGCCCC 838129
```

```
RESULT 10
US-08-804-227C-7/c
; Sequence 7, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44377 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  350..14002
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14046..20036
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20110..31284
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31329..36071
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36155..41830
US-08-804-227C-7

Alignment Scores:
Pred. No.:              5.22e-21      Length:        44377
Score:                  670.00        Matches:       665
Percent Similarity:     31.22%        Conservative:  282
Best Local Similarity:  21.93%        Mismatches:    1121
Query Match:            5.07%         Indels:        980
DB:                     2             Gaps:          149

US-09-522-753-5 (1-2517) x US-08-804-227C-7 (1-44377)

Qy          9 AlaGlnThrTrpArgAlaThrGluProArgTyrProProHisSerLeuSerTyrProVal 28
              |||:::||||||   |||      |||            |||            |||
Db       9254 GCCCGTACGTGGCCAGCAGCGCACCCGCCTCGATCGGATCACCCAGCGCCGTCCCCGTAC 9195

Qy         29 GlnIleAlaArgThrHisThrAspValGlyLeuLeuGluTyrGlnHisHisSerArgAsp 48
              :::            |||
Db       9194 CGTGCGCCTCCACCACAT------------------------------------------ 9177

Qy         49 TyrAlaSerHisLeuSerProGlySerIleIleGlnProGlnArgArgArgProSerLeu 68
                                                  |||:::||||||   |||:::
Db       9176 ------------------------------------CCACGTCGGCGGGCACCAGCCCCG 9153

Qy         69 LeuSerGluPheGlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSer 88
                                |||:::      :::         |||      |||   :::
Db       9152 CATCAGCCAGCGCCTCACGGATCACCCGCTGCTGCGCCGGACCACTCGGCGCCGTCAGCC 9093

Qy         89 HisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIleGluSerLysArgPro 108
              |||                                             |||   ||||||
Db       9092 CATTGC------------------------------------------TCGCACCGTCCT 9075

Qy        109 ArgLeuGluLeuLeuProAsp------------------------ProLeuLeuArgProSer 121
                 |||      ||||||                        ||||||   |||   |||
Db       9074 GA-TTGACCGCACTGCCCCGCACCACCGCCAGCACCCGATGACCGTTGCGGCGCGCGTCG 9016

Qy        122 ProLeuLeuAlaThrGlyGln---ProAlaGlySerGluAspLeuThrLysAspArgSer 140
                       :::|||         |||      |||                  |||:::
Db       9015 GAAAGCCGCTCCACCAGCAGCACACCCACACCCTCCGCCCACGCCGTCCCGTC-CGCACC 8957

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProPro------------------ 154
              |||      :::|||   |||               ||||||
Db       8956 CTCCGCGAACGCCTTGCACCGCCCGTCCGCCGCCAGCCCCCGCTGCCGCGAGAACTCCAC 8897

Qy        155 ---------------------HisThrAspPro------------GluLeuGluLeuVal 163
                                   |||      |||            ::::::   |||
Db       8896 GAACACCGTCGGCGTCGCCATCACCGTCACCCCACCGGCCAACGCCAGATCACACTCACC 8837

Qy        164 ProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGluIle 183
```

```
             ||||||            :::   |||   :::   |||   |||      |||
Db      8836 CCGCCG------------CAACGACTGCACCGCCAGATGCAACGCCACCAACGACGACGA 8789

Qy       184 ThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu---Glu 202
             |||                     ::::::   ::::::::::::   |||:::   :::
Db      8788 ACA-------------------CGCCGTGTCCACCGTCAGCGCCGGCCCCTCCAACCCCAA 8747

Qy       203 GluAlaAlaLysPro-ProGlu---------ProGluLysProValSer------ProPr 217
                         ||| ||||||         |||      |||||||||      |||||
Db      8746 CGCATAGGCAACCCGACCGGAGATCACACTGCCGAGGGTCCCCGTGAGCACATGCCCGCC 8687

Qy       217 oProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLy 237
             |   :::   :::         |||   |||   ::::::         :::      ::
Db      8686 GTAGGTGTCGGCGGCGTCCCCCAGCTGGGTCGC-GTAGTCCTGCTGGGAGATGCCGCTGA 8628

Qy       237 s---------AlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLe 254
             :               ||||||   |||:::            ||||||         ||
Db      8627 AGACACCGGTCCGGGTGCCGCGCAGCGAGTCCGGAACGATGCCGGCCCGCTCCAGCGCCT 8568

Qy       254 uPro------LeuTyrAsnGlnPro--------------------SerAspThrArgGl 265
             ||||      |||      ||||||                              |||::
Db      8567 CCCACGACGTCTCCAGCAGCAACCGCTGCTGCGGGTCCATGGCCGTGGCCTCGCGCGGCG 8508

Qy       265 nTyrHisGluAsnIleLysIle------AsnGlnAlaMetArgLysLysLeuIleLeuTy 283
             :|||   |||   :::::::::         :::      ::::::            ||
Db      8507 ATATCCCGAAGAACTCCGCGTCGAACTCGGCCGCGGAACGCAGGAACCCGCCTTCCCGTA 8448

Qy       283 rPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAs 303
             |               |||   |||               :::      |||      ||
Db      8447 CATAGCTCTTGCCCGGCACCCCCGG------------CTCCGGGTCGTACAGCCCCTCGA 8400

Qy       303 pGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgAr 323
             |                     ::::::   ||||||:::
Db      8399 CGTCCCAGCCACGGTCGGCCGGGAACGGGGAGAGCGTGTGCCCGTCACCGGCCAC----- 8345

Qy       323 gAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLysGl 343
                                                  |||   |||::::::   :::::
Db      8344 -------------------------------------CAGGTCCCACAGGTCGTCCGGCG 8322

Qy       343 nArgGluLeuGlnGluArgMetGlnSerArg----------------------------- 353
             :|||            |||:::   :::|||
Db      8321 ACGCCAC---ACCGCCCGGGTAACGGCACGCCATGCCCACGATCGCGATCGGCTCACGCT 8265

Qy       354 -ValGly------GlnArgGlySerGly---------------LeuSerMetSerAlaAl 366
              ||||||      :::||||||   |||               |||::::::
Db      8264 CGTCGGCCATGACCGGCGTGGGGGCGGCCGCTGCCGGCTCCGCCTCGCTGTCGGGACCGA 8205

Qy       366 aArg------------------SerGluHisGluValSerGluIleIleAspGlyLeuSe 380
              :::                     |||   |||         :::::::::
Db      8204 ACAGTTCGTCGTGCAGGTGGGAGGCGAGCGAGAGGGGGGTCGGGTGGTCGAA-------- 8153

Qy       380 rGluGlnGluAsnLeuGluLysGlnMetArgGlnLeu----------------------- 392
                             |||:::::::::||||||:::
Db      8152 ----------------GAGCAGCGTGTTCGGCAGGTTCAGGCCGGTGGCAGCGGTCAGGC 8109

Qy       393 -------------------------AlaValIlePro------ProMetLeuTyrAspAl 402
                                         |||   |||      |||
Db      8108 GGTTACGCAGTTCGAGAACGGTGAGGGAGTCGAACCCGGCGGACCG------------GA 8061

Qy       402 aAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro------Me 420
                 |||   |||:::      :::         |||:::      ||||||      ::
Db      8060 AGGCCAGCTCCGGGTCGACCGCGTCGGC---CGAGGCGTGGCGCAGGATCCCGGCGGCCT 8004

Qy       420 tLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPh 440
             ::::   :::|||                           ||||||            ||
Db      8003 CCGAGCGCACAAG---------------------------CGTGAGCAGAGCACGACGGCGTT 7968

Qy       440 eArgGluLys-------------------------------------------------- 443
             ||||   :::
Db      7967 CCGCCGGCGAGGGGCCGACGGAGAGCCACGGCCGGCCCTCGCCCCCGTCCTGTCCGCTCC 7908

Qy       444 ----------PheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu-- 459
                          :::      |||:::   :::|||         :::   :::
```

```
Db      7907 CGCCCGGTGCGCCGTCGCCCGGCCGCGTCCCTGGGGTGAGAGCGGCACGGTCGTCCGCGG 7848

Qy       460 -------------ArgLysThr-ValAlaGluCysValLeuTyrTyrTyrLeuThrLysL 475
                          |||:::|||    ||||||
Db      7847 CATGGCCGGCACGCGCCGAACGGGCGGCGGAG---------------------------- 7816

Qy       475 ysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnG 495
                     :::            |||||||||         |||      |||      |
Db      7815 --ACGCCGTCGAAGAGCGACCAGGTACGGCGGGTGACCGAACGCCCGACGAAGGTCTCCC 7758

Qy       495 lnGlnGlnGlnGlnGlnGlnGlnGlnGln------------------GlnGlnGlnG 509
             ||                  ::::::                              :::
Db      7757 AGTCCACGTCGGCGACGAGCAGGGAGCTCTCGCCGGCCGCGTGCGTCCCTGCGGCAGAGG 7698

Qy       509 lnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluA 529
                  |||   |||   :::|||                  :::         :::
Db      7697 TCCCGCCGGCGCCCACCACGTCGTCGAGGACGTCGAGCGCGCGCTCGGGGTCCAGGGCCG 7638

Qy       529 laGluLysGluGluGluLys---------------ProGlu------------------- 537
                              |||               ||||||
Db      7637 TGACGCCGTCCCGTTCGAAGGACTTCGCCGCCGTTCCGGAGGCCATTCCGGCACCGGCCC 7578

Qy       538 --------------ValGluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspT 553
                           :::|||      :::      |||   :::   :::
Db      7577 ACGGCCCCCAGGCGATCGAGGCGGCACGGGCTCCCCTGGCCCGGCGCCGCTGCGCCAGGG 7518

Qy       553 hrSerGly-----------------------------------------------GluA 557
               |||
Db      7517 CGTCGAGCGCGGCGTTGGCGGCCGCGTAGACGGCCTGCCCTCCGTTGCCCCACACGCCGA 7458

Qy       557 spAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyA 577
                  :::||||||         ||||||   ||||||         :::|||
Db      7457 CGATGGAGGAGAAGAGCACGAAGGCGTCGAGCGGAAGGTCCGCGGTCAGCTCGTCCAGGT 7398

Qy       577 rgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleT 597
               |||                                                |||:::|
Db      7397 GACGG---------------------------------------------GCGGCGGTGA 7383

Qy       597 hrProGlnGlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrG 617
             ||      :::|||                  |||::::::|||||||||      |||
Db      7382 CCTTGTCCCGCAGTACGGCGTCCACCCGGTCGGAGGTGAGCGAGTCGAGC------ACAC 7329

Qy       617 luGluGluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaI 637
                              |||                     ||||||:::
Db      7328 CGTCGTCGAGCACACCTGCCGCGTGCAGTACGGCGACGAGGGGCCGCTCCTCGGGGAGGC 7269

Qy       637 leAlaArgMetValGlySer---------LysThrValSerGlnCysLysAsnPheTyrP 654
               ||||||         |||                  ||||||            :::|
Db      7268 CCGCGAGGAGCGCGGCCAGTGCGTCCCGGTCGGCGACGTCGCAGGCCTCGACGGC-CACT 7210

Qy       654 heAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGln--------------- 668
             ||      ::::::||||||   |||||||||
Db      7209 TCGGAACCGAGCGCCGTCAGCTCCTCGACGAGTTCGCCCACTCCGGGGCCGGCCGCGCCG 7150

Qy       669 --------HisLysLeuLysMetGluLysGluArgAsnAlaArg-ArgLysLysLysLys 685
                     |||:::         |||:::   |||   |||||| |||         :::
Db      7149 CGCCGGCCCACCAG---------GAGCAGACGCGCGGCGCCCGACCGGGCGAGCCTGCGG 7099

Qy       686 AlaProAlaAlaAlaSerGluGluAlaAlaPheProProValValGlu-----AspGlu 703
             |||      |||            |||      |||||||||:::            |||
Db      7098 GCGACCTGCGCA------CCCAGGGCACCGGTGCCCCCGGTGATCAGTACGGTGCCGGAG 7045

Qy       704 GluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGluAlaGluAla 723
             :::      :::||||||   :::   :::|||                  :::
Db      7044 GACCACCGGTCGTCCGGCACGACCACGTCCGAGGTGTCGTTCCGTGCGTCCTCGCCGACG 6985

Qy       724 LeuHisAlaSerGlyAsnGluValProArgGlyGluCys---SerGlyProAlaThrVal 742
                   |||                  |||   |||      :::|||
Db      6984 GCGGTGGCGGGGGCCGTGGTGACGGTGCGAGCGGAGGCCGGGACGGGGACGGCGCGGG- 6926

Qy       743 AsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLysAspThr 762
             |||:::::::::|||   ::::::         |||:::   |||
Db      6925 AACCACACGGCGGACATGGACGCC---GGAGGGCCGTACCGCGAACTGGTTCTCTCCCCC 6869
```

```
Qy         763 GlyGlnAsnGlyProLys---------------ProProAlaThrLeuGlyAlaAspGly 777
               |||:::      |||:::               |||   ||||||            |||
Db        6868 GGCGAGCA-GCGCCCCGACGAACGTACGGAGCACCGCGGGCGACGCACTCCCGGGCAGGT 6810

Qy         778 ProProProGlyProProThrProPro---------ArgArgThrSerArgAlaProIle 794
               ||||||      |||         |||         ||||||      |||   |||
Db        6809 CCGCCAAGCCGCCCCAGCGCCGCCCCAGTTCGAGCGCGGCGACCTGTCCGAGTCCCCACA 6750

Qy         795 GluProThrProAlaSerGluAlaThr---GlyAlaProThrProProPro--------- 810
                        |||      :::   :::      |||      |||||||||
Db        6749 GCTGAGCGCCCCCGACGGAAGGGGTCTCGTCCGGCCGCACGGCCGCCGCCTCACGGGTGG 6690

Qy         811 ---------------------AlaProProSerProSer--------------------- 816
                                       ||||||   |||:::
Db        6689 TGATCCAGAGCGGGGCTTCGGGTACCGCCTCGCCCAGCGCCTGGACGAGGGCGACGGCGT 6630

Qy         817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
                  |||   |||      |||            |||               |||   |||
Db        6629 CGGCCGCGTCCGGCTCGGCCGGGGCGAGCCACAGAACT-CGGGTGCTGCGGGCGGCACCG 6571

Qy         837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
                                 :::   |||   |||:::            |||   ||||||
Db        6570 TCACCGGCGGCCCGCAACAGGTCCCCGGCTGCCTCCCGG------CCTGTACCGACCGGC 6517

Qy         857 LysAlaGluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLys 876
                  |||         |||   :::               |||
Db        6516 ACGGCG------ACGGTGTGGACCTCGGCCCCGCGCACGGCGAGTTCACGTGCGATCGCG 6463

Qy         877 GlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLys 896
               |||         |||:::|||      ::::::   ||||||      :::
Db        6462 GGGGCGAGTCCGGAGTCGGCC---GTGTCCTCGTCCGGGGCCAGAAGCAGCACATGACCG 6406

Qy         897 GluGly---GlySerGlyArgAlaThrThrAla--------------------------- 906
                  |||   |||      |||:::|||||||||
Db        6405 GCAGGCCCGGGGTTGCCGCGGTCGACGACGGCCGGCTGCCAAGCGACGCGGTAGCGGATC 6346

Qy         907 ------------------------------------LysSerSerGlyAlaProGlnAsp 914
                                                      |||:::|||   |||
Db        6345 CCGGCCGTCGTCGACCGTCCCCGCTCAGTGGCCCGAGCGTCCGCCGGCGAGCCGGCGGGC 6286

Qy         915 Ser--------------------------------------------------------- 915
               |||
Db        6285 TCGACGCCGGTCACGGCGGTGCGAACGGGGTCCAGCCAGTAGCGGTCGCGTTGGAAGGGG 6226

Qy         916 ------------------------------------------AspSerSerAlaThrCys 921
                                                            :::   |||:::
Db        6225 TAGGTGGGGAGGTCGATGCGCCGGCCGGTGCTCCCGGTGTGGAGTGCCGTGGCGTCCAGG 6166

Qy         922 SerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArg 941
                  |||      |||:::                     |||         |||   |||
Db        6165 CCGGCGTCCCGGGTGAACACCGTCGCCAGCGCCGCCTCGAACACCTCCCGCTCCGCACGG 6106

Qy         942 ProSerLeuLeu---------ThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
               |||      :::         |||   |||      |||            ||||||
Db        6105 CCCCGCCGCATCGCCGGCACCACCACCACATCATCACCGGCCCCCAGGCACTCACCCGC- 6047

Qy         959 LysPro------------------LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAla 972
                  |||                     |||   |||
Db        6046 CATCCCCGTCAGCACCCCATGCGGACCCACCTCCACCAACGTCCCCACACCCAGACCACG 5987

Qy         973 IleProProIleGlnValThrLysValHisGluProProArgGluAspAla------Ala 990
                     |||      :::   :::   |||   ||||||      ||||||
Db        5986 CACCACCCCCACCCCATCCGCGAAACGCACCG-CCTCCCGCGCATGACGCACCCAATACC 5928

Qy         991 ProThrLysProAlaProProAlaPro------------------------ProProPro 1002
               |||      |||   ||||||   |||                        |||||||||
Db        5927 CCGGATCCCCCAACTCCCCCGAACCCACCACCCCACCCGACACACCCGACACCACCACCA 5868

Qy        1003 GlnAsnLeuGlnProGlu------SerAspAlaProGlnGlnProGlySerSerProArg 1020
                           |||:::               ||||||   ||||||:::   |||
Db        5867 CACCCGGCCGCACCCGACCGAACTCCAACGACTCCACAACCCCCCGGAACTCCCCCA--- 5811

Qy        1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
                           :::|||||||||            |||      |||
Db        5810 ------------ACACCGGCTCCATCAACACCGAATGAAACCCATGCG------ACACAT 5769

Qy        1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
                        |||   ||||||         |||            |||   |||
Db        5768 CCAACCGCCGGCACTCCACCCCATCCCCCATCAGCGAGGCAACCACCGACTCCA------ 5715

Qy        1061 GluValIleLysAlaSerProHisAlaPro---AspProSerAlaPheSerTyrAlaPro 1079
                           |||   ||||||   |||      |||         :::::::::|||
Db        5714 ------------GCACACCCACATCACCCGACAACACCACCGACCGCGGCCCATTCACCG 5667

Qy        1080 Pro-Gly---HisProLeuProLeuGlyLeu---------------HisAspThrAlaAr 1093
               |||       ||||||||||||      |||               |||         ||
Db        5666 CCGCAACCGACACCCACTCCCCCAACCCCTCAACAACCCCCCGCACCACCGACTCCGACG 5607

Qy        1093 gProValLeuProArgProProThrIleSerAsnPro----------------------- 1105
               ||||      |||   |||||||||         |||
Db        5606 CCCCCACCGACCACATCCCCCCACCCACCGGCAACCCACCCATCAACCCACCCCGCGCCA 5547

Qy        1106 -------ProProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGl 1123
                      |||      |||:::      :::||||||      :::
Db        5546 CCACCAACCGCACCGCATCACCCAACGACAACACCCCCGCCACATACGCAGCAGC----- 5492

Qy        1123 yAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAlaLysAl 1143
                                              |||:::|||:::
Db        5491 -------------------------------CACCTCCCCCACCG-AATGACCCAACACC 5464

Qy        1143 aProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLysLeuAlaPr 1163
                |||      |||               |||   |||      |||            ||
Db        5463 ACCGACACCTCCA------------CACCCCGAGCCTCCAACGCCCGGAACAACGCCACC 5416

Qy        1163 oPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro-------- 1180
               |   :::         |||   :::   :::|||               |||
Db        5415 TCCAACGCAAACAGCGCAGGCTGAGCAAACTCCGTCCGACCCAAC-AACCCACCCACAGA 5357

Qy        1181 -------ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyTh 1198
                      ||||||            |||            ||||||
Db        5356 ACCGGAACCAGAACCGACCCCAGCACCCGCACCGGCACCCGCATCAGCCCCGGCACCCGC 5297

Qy        1198 rAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPr 1218
                :::      :::   |||      :::               |||||||||
Db        5296 GTCCACGTCGACGTCGCCGAACATCACATCCCGCAACGACCGACCATCCACCTCCCCCAC 5237

Qy        1218 oSerAspSer------------------------------------AlaIleThrTyrAr 1226
                   |||:::                                    |||         ||
Db        5236 CATCGACAACACCTCATCCAGCACCTCCGCAAACACCCCACCCCCCGCATACAACCCACG 5177

Qy        1226 gGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIl 1246
               |         ||||||      |||      |||            |||   |||
Db        5176 CCCCATACCAACCCACTGCGCACCCTGACCCGTG------AACAACACCCCCACC----- 5128

Qy        1246 eIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGl 1266
                                  :::|||   ||||||         |||:::||||||
Db        5127 -------CCACCCGGCACAGCCCGCCCGGACCGCACACCCCCAGACACCCTCCCGGCAGC 5075

Qy        1266 yHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerVa 1286
                :::
Db        5074 AAAC-------------------------------------------------------- 5071

Qy        1286 lThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAl 1306
                                           ||||||      |||||||||   |||.  ||
Db        5070 ----------CCACCAAGAGCAGCGACGAGCTCCCCACGCCCACCACACGGCACCACCGC 5021

Qy        1306 aProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSe 1326
               |                              |||      :::   :::         ::
Db        5020 A------------------------------CGGTGCTCGAAAAGAGTCCGCGTCGAAGC 4991

Qy        1326 rIleGluGlyLeuMetGlyArgAlaIleProPro-------------------------- 1337
               :   |||            |||         |||
Db        4990 CAGCGAATAAGCCACATCCCGCGGATCCGCACCCACCCGCTCCACGTACTCACCCAACCG 4931

Qy        1338 ----------GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySe 1354
```

```
                           |||::::::|||            :::   :::         |||::
Db       4930 CACCGCCTGCGCCCGCAACGCCCCCTCCGACCGCGGACAACAACAGCGGCGACGGAGC 4871

Qy       1354 rIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGl 1374
              :            :::|||||||||                  |||   |||
Db       4870 AGCCGCGTCCGCGGTACCGCGGTCAGGAGCAGCGACGCCGTC-GATGTCCTCGACCCACT 4812

Qy       1374 uAlaLysLeuLeu------------------------------------LysArgGluGl 1382
                  :::||||||                                    :::|||
Db       4811 CCGGCGCCTCCTCGATGATCAGATGCGCGTTCGTCCCACTGATCCCGAACGACGACACAG 4752

Qy       1382 yThrPro---------------ProProProProProSerArg---------------As 1392
                  |||               ||||||   ||||||:::|||               ||
Db       4751 CCGCCCGGCGCGGGCGGTCCGACCGCCGGGGCCACCGGCGCGCCTCTGCCAGCAGGGCGA 4692

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              |                              ||||||:::::::::|||         ::
Db       4691 CGGTGCCCGACGACCAGTCGACGTGCGTGCTGGGCCGGTCCGCGTGAAGTGTCCGCGGCA 4632

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
              :      |||         |||                                 :::
Db       4631 GCGTGCCGTGCTCCAGGGCAAG---------------------------------CAGCA 4605

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
               |||   ||||||                  ||||||            |||   :::
Db       4604 TCTTGACCACACC------------CGCGGCCCCGC------------GGCGGCCTG-- 4571

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                                          |||   :::      |||   ::::::   ::
Db       4570 ----------------------------GGTGTGACCGACGTTGGACTTCAGCGACCCGA 4539

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              :   :::   :::                                 |||||||||
Db       4538 GCCACAAGGGGTC-------------------------------GCCCACCCGCTC----- 4514

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                           |||      :::||||||      :::         :::   |||||
Db       4513 ----ACGCCCGTACGTGGCCAGCAGCGCGCTCGCCTCGATCGGATCGCCCAGCGCCGTCC 4458

Qy       1512 oGlyThrAla------------------------------------------------Se 1516
              |   |||
Db       4457 CCGTACCGTGCGCCTCCACCGCGTCCACGTCACCCGGCGCGAGGCCGGCGTCGGCCAGGG 4398

Qy       1516 rSerSerGlyGlySerIleAlaArgGlyAla--------------------------Pr 1527
               ||||||      ::::::   |||||||||
Db       4397 CAGCTCGGATGACGCCCTCCTGCGCGGAGCGGTTGGGAGCGGTCAGGCCGTTGCTGGCAC 4338

Qy       1527 oValIleVal--ProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHis 1546
               |||:::|||   |||            ||||||   :::|||
Db       4337 CGTCCTGGTTGACCGCACTGCCCCGCACCACGCCCAGCACCC------------------ 4296

Qy       1547 GlyAlaProPheAlaGlyHisLeuProArgGlySerProValThrMetArg--------- 1563
              |||   |||   ||||||         |||   :::||||||   :::|||
Db       4295 GGTGTCCCGCACGCCGGGCGT---CGGAGAGCCGCTCCAGTGCCAGTACGGCGACACCCT 4239

Qy       1564 ------GluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGln 1581
                       |||      |||   :::      :::      |||   :::      :::
Db       4238 CGGACCAGGCCGTACCGTCGGCATCGGCGGAGAACGGCATGCAGCGGCCGTCGGGGGCGA 4179

Qy       1582 AspArgLysLeuThrSerThrProArg------GluIleAlaLysSerProHisSerThr 1599
                 |||            |||||||||                  ||||||
Db       4178 GCCCGCGCTGCCGCGAGAACTCCACGAACATGCCGGGCCCCGACATCACCGTGGCGCCGC 4119

Qy       1600 ValProGluHisHisProHisProIleSerPro--------------------------- 1610
                 ||||||      ||||||      ||||||
Db       4118 CCGCCAGAGCC-AGCCCGCAC------TCGCCCCGGCGCAGCGCCTGCACGGCCAGGTGC 4066

Qy       1611 ---------------TyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArg 1625
                                ||||||   :::      ||||||
Db       4065 AGCGCCACCAGGGACGAGGAGCACGCGGTGTCGACGGTGAGC------------------ 4024

Qy       1626 SerHisIleProLeuAlaPheAspProThrSerIleProArg------GlyIleProLeu 1643
                             |||      |||      :::||||||         |||
Db       4023 ---------------GCCGGTCCGCCG------GTGCCGAGGGTGTAGGCGATCCGGCCG 3985

Qy       1644 AspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHis 1663
              :::|||      |||                  :::   |||         ||||||
Db       3984 GAGGCCACGCTCGCG-----------------GTGGTGCCGGTGAGCAGGTACCCCTCG 3943

Qy       1664 LeuTyrPro------------------------ProTyrLeuIleArgGlyTyrPro--- 1674
                 ||||||                        ||||||               |||
Db       3942 ---TATCCGTCCGCTCCCTCGTGAAGCCTCGGCCCGTACTCCGGTGCGGTGGCGCCGACG 3886

Qy       1675 AspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGln 1694
              :::|||   |||         |||            :::         :::|||:::
Db       3885 AACACTCCGGCGTCGCCGCCGCGCAGTGACTCGGGGAGCACTCCGGCGTCCTCGAGGGCC 3826

Qy       1695 MetHisHisAsnThrAlaThr---------------------AlaMetAlaGlnArgAla 1707
                 |||      :::::::::                     |||:::|||   |||
Db       3825 TCCCACGACGCTTCCAGCAGCAAGCGCTGCTGCGGGTCCATGGCCGTGGCCTCGCGCGGC 3766

Qy       1708 AspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGly 1727
              |||:::               |||:::   |||      |||   |||   ||||||
Db       3765 GATATC---------------CCGAAGAACTCC------GCGTCGAACTCGGCCGCC--- 3730

Qy       1728 ProArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThr 1747
                             :::            |||            ::::::|||
Db       3729 ---------------GAGTGCAGGAACCCGCCCTCCCGCACATAGGTCTTGCCGGGCGTC 3685

Qy       1748 ProGly------------ThrProAlaThrAlaMetAspArgLeuAla---------Tyr 1760
              ||||||            :::|||:::   :::      |||
Db       3684 CCGGGCTCGGGGTCGTACAGCCCGTCGAGGTCCCAGCCGCGGTCGGTGGGGAAGGACGAG 3625

Qy       1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
              :::   :::||||||         ||||||:::   ::::::         :::|||
Db       3624 ATGGCGTCGGCGCCGGACTCCAGCAGCTCCCAGAGAGCGGCGCGAGTCGACGCCGCCG 3565

Qy       1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
              |||            |||   ||||||            |||         |||
Db       3564 GGGAGCCGGCAGCCGACA---CCCACGATGGCGATCGGGTCGTCGTCGTCCCGGCTGCCG 3508

Qy       1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
                 |||      |||               ::::::|||                  |||
Db       3507 CCCCGGGCGTCCGACACCCCTACCGCGGCCGACGAGAGCACTCCGGGCCCACCGGCAACC 3448

Qy       1821 ValGluHisAlaPro------------IleTrpArgProGlyThrGluGlnSerSerGly 1836
                          |||                           |||   :::      |||
Db       3447 GAGTCCTTCGGCCCGCCCGCGGCTTCCACGATGTGCTCGGCGACGGCACGCGGGGTGGGC 3388

Qy       1837 SerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHis 1856
                 |||   :::|||            ||||||         |||   |||   ||||||
Db       3387 AGGTCGTAGACCAGCGTGGCGGCGAGGGGGAGCCCCGAGGCGCCCTTCAGCAGGTCGCAC 3328

Qy       1857 -AlaHisGlnHisSerProIleSerPro------------------------ArgThrGl 1868
               |||         ||||||   ||||||                        |||   ::
Db       3327 AGCTCGACCGCGGTCACCG---AGTCCATGCCCTGGGAGGTGAAGGTGCGGTCGAGCGCG 3271

Qy       1868 nAspAlaLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleTh 1888
              :            :::|||
Db       3270 ACCTCGCCGGGATCGTCGT------------------------------------------ 3252

Qy       1888 rAlaValGluProSerLysProThrValLeuArgSer------------------ThrSe 1902
                        |||   :::|||         |||
Db       3251 -------GGCCCAGGACGGCCGCGGTGCTCTCGCGGACCAGGCGGACGGCATCCTCGGGG 3199

Qy       1902 rThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGl 1922
               ||||||:::||||||   |||:::         ||||||         ||||||   ||
Db       3198 GACATCGGCTCCGGTGACGCCGAGT---------CCGCCTCGTCGTCCGTGTCCCTGTGG 3148

Qy       1922 yGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGluAl 1942
              |   :::                                          |||:::   ::
Db       3147 GACGTCC------------------------------------------CCTCGGTGCTC 3130

Qy       1942 aProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLy 1962
              :|||         |||         ||||||            |||   :::   :::
Db       3129 TCCGGGTCCGGGGAGACGCTGTGCCCCACGG------------CATCGGTGGCGGTCACA 3082
```

```
Qy       1962 sProProAlaArg---------------------------------SerGlyLeu----- 1969
               ||||||   |||                                 ||||||
Db       3081 CCCGCCCCGGCGTCGGCCGTGACCGGCCGGCTCAGCGCGGGCGACCAGTGGGTACGGCGT 3022

Qy       1969 ------------------------------------------------------------ 1969

Db       3021 TGGAAGGGGTAGGTGGGGAGGTCGATGCGCCGGCCGGTGCTCCCGGTGTGGAGTGCCGTG 2962

Qy       1970 -GluProAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSe 1989
                  |||         |||               |||   |||         |||   ::
Db       2961 GCGTCCAGGCCGGCGTCCCGGGTGAACACCGTCGCCAGCGCCGCCTCGAACACCTCCCGC 2902

Qy       1989 rGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerPr 2009
              :   |||         |||   :::||||||            |||||||||      ||
Db       2901 TCCGCACGGCCCC---GCCGCATCGCCGGCACCA------CCACCACATCATCACCGGCC 2851

Qy       2009 o---------AspProProAlaProProAlaSerAlaSerAsp----------------- 2020
              |            ||||||:::|||   |||      :::|||
Db       2850 CCCAGGCACTCACCCGCCATCCCCGTCAGCACCCCATGCGGACCCACCTCCACCAACGTC 2791

Qy       2021 -ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
               ||||||   :::               |||   |||   :::            |||::
Db       2790 CCCACACCCAGACCACGCACCACCCCCACCCCA-TCCGCGAAACGCACCGCCTCCCGCGC 2732

Qy       2040 r------LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerPr 2058
              :            |||   ||||||      ||||||         ||||||   :::||
Db       2731 ATGACGCACCCAATACCCCGGATCCCCCAACTCCCCC---------GAACCCACCACCCC 2681

Qy       2058 oValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAs 2078
              |      :::|||      |||         |||         |||      ||||||
Db       2680 ACCCGACACACCCGACACCACCACCACACCCGGCCG-------CACCCGACCGAACTCCA 2628

Qy       2078 pLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGl 2098
               :::   |||                                 |||   :::|||
Db       2627 ACGACTCCAC------------------------------AACCCCCCGGAACTCCCCCA 2598

Qy       2098 yGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerPr 2118
                        |||         ::::::|||:::         ||||||         ||
Db       2597 ACACCGGCTCCATCAACACCGAATGAAACCCATGCGACACATCCAACCG---------CC 2547

Qy       2118 oLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHi 2138
              |   |||         |||      :::|||:::   |||:::            |||||
Db       2546 GGCACTCCACCCCATCCCCCATCAGCGAGGCAACCACCGACTC-----------CAGCA 2499

Qy       2138 sIleSerGluValIleThrGlnAspTyrThrArgHisHisPro------GlnGlnLeuSe 2156
              |            ||||||:::   :::   |||      |||      ::::::
Db       2498 CACCCAC------ATCACCCGACAACACCACCGACCGCGGCCCATTCACCGCCGCAACCG 2445

Qy       2156 rAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLe 2176
                  |||||||||   ||||||   :::   |||                        ||
Db       2444 ACACCCACTCCCCCAACCCCTCAACAACCCCCCGCACCACCGA---------------CT 2400

Qy       2176 uArgArgProProSerAspLeuTyrLeuPro---ProProAspHis-------------- 2190
              ||||||||||            :::   |||   ||||||   |||
Db       2399 CCGACGCCCC------------CACCGACCACATCCCCCACCCACCGGCAACCCACCCA 2352

Qy       2191 ----GlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluPr 2209
                        |||            ||||||            ::::::   |||
Db       2351 TCAACCCACCCCGCGCCACCACCAACCGCACCGCATCACCCAACGACAACACCCCCGCCA 2292

Qy       2209 oAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGl 2229
                  ::::::|||   |||                                    |||
Db       2291 CATACGCAGCAGCCACCTC-------------------------------------CCCCA 2268

Qy       2229 uGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeu----------- 2245
                  ||||||:::      |||   |||            |||   |||
Db       2267 CCGAATGACCCAA------CACCACCGACACCTCCACACCCCGAGCCTCCAACGCCCGGA 2214

Qy       2245 ------------------------------------------------------------ 2245

Db       2213 ACAACGCCACCTCCAACGCAAACAACGCAGGCTGAGCAAACTCCGTCCGACCCAACAACC 2154

Qy       2246 ----TyrArgAspGlyGluGlnThrGluProSer---ArgMetGlySerLys-Ser---- 2261
                  :::|||   |||   :::|||:::||||||   |||   |||:::::: |||
Db       2153 CACCCACAGAACCGGAACCAGAACCGACCCCAGCACCCGCACCGGCACCCGCATCAGCCC 2094

Qy       2262 -----ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnS 2280
                   |||   :::||||||      |||               :::|||      |||
Db       2093 CGGCACCCGCGTCCACGTCGACGTCGCCG------------AACATCACATCCCGCAACG 2046

Qy       2280 erAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgA 2300
                            |||         :::                  :::   |||
Db       2045 ACCGACCATCCACCTCCCCCACCATCGACAACACCTCATCCAGCACCTCCGCAAACACCC 1986

Qy       2300 snGluProGluTyrAsnIleSerGlnProGlyThrGlu---------------IlePheA 2315
                   |||   ||||||               |||                  :::   |
Db       1985 CACCCCCCGCATACAACCCACGCCCCATACCAACCCACTGCGCACCCTGACCCGTGAACA 1926

Qy       2315 snMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluH 2335
              ||   |||         ||||||
Db       1925 ACACCCCCACCCCACCCGGCACAGCC---------------------------------- 1900

Qy       2335 isAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrA 2355

Db       1900 ------------------------------------------------------------ 1900

Qy       2355 spGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerA 2375
                :::   :::   :::||||||   :::      |||   ||||||      |||:::
Db       1899 --CGCCCGGACCGCACACCCCCAGACACCCTCCCGGCAGCAAACCCACCAAGAGCAGCGA 1842

Qy       2375 laSerLeuPro------------------------------------------------- 2378
                |||   |||
Db       1841 CGAGCTCCCCACGCCCACCACACGGCACCACCGCACGGTGCTCGAAAAGAGTCCGCGTCG 1782

Qy       2379 --AlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProG 2398
                |||:::         |||:::                     :::      ||||||
Db       1781 AAGCCAGCGAATAAGCCACATCCCGCGGATCCGCGCCCACCCGCTCCACGTACTCACCCA 1722

Qy       2398 lyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProA 2418
                            |||:::            ||||||   |||   |||   :::   :
Db       1721 ACCGCACCGCCTGCGCCCGC------AACGCCCCCTCCGACCGC---GCGGACAACAACA 1671

Qy       2418 laProGlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyA 2438
              ::   |||   :::      :::               :::   |||      ||||||
Db       1670 GCGGCGGGACGTCCGGAAGCGAGGCGGCGACGGAGACGGCCGGGGTGGCCGGGAGGGCA 1611

Qy       2438 spCysAsnArgArgThr---------------ProLeuThrAsnArgValTrpGluAspA 2453
                |||:::|||   |||               |||:::   |||         |||
Db       1610 GGTGTTCCAGGACCACATGGCAGTTGGTGCCACCGATGCCGAAC---------GAACTGA 1560

Qy       2453 rgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnA 2473
                |||:::|||   |||:::   |||
Db       1559 CACCCGCCAGCAGGGGTGCGCCGCCC---------------------------------- 1534

Qy       2473 laGlyValMetAlaSerProProProPro-GlyLeuPro---AlaGlySerGlyProLeu 2491
                |||         |||                   |||   ||||||:::||||||
Db       1533 --GGGGCGAGCGGCAGCTCGGCGGCAGCGGTCTGCACCTTCAGCCGGAGCTGGTCCAGCG 1476

Qy       2492 AlaGlyProHisHisAlaTrp 2498
                    |||         |||
Db       1475 GGATGGCCGGAGAGGGGGTGG 1455
```

```
RESULT 11
US-08-804-198-1/c
; Sequence 1, Application US/08804198
; Patent No. 5945320
;  GENERAL INFORMATION:
;    APPLICANT:  Burgett, Stanley G.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rao, Nagaraja R.
;    APPLICANT:  Richardson, Mark A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    TITLE OF INVENTION:  PLATENOLIDE SYNTHASE GENE
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  PAUL R. CANTRELL  1138
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh 7.0
;      SOFTWARE:  Microsoft Word 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,198
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CANTRELL, PAUL R.
;      REGISTRATION NUMBER:  36,470
;      REFERENCE/DOCKET NUMBER:  P9113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3885
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44377 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  350..14002
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14046..20036
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20110..31284
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31329..36071
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36155..41830
US-08-804-198-1

Alignment Scores:
Pred. No.:              5.22e-21      Length:        44377
Score:                  670.00        Matches:       665
Percent Similarity:     31.22%        Conservative:  282
Best Local Similarity:  21.93%        Mismatches:    1121
Query Match:            5.07%         Indels:        980
DB:                     2             Gaps:          149

US-09-522-753-5 (1-2517) x US-08-804-198-1 (1-44377)

Qy          9 AlaGlnThrTrpArgAlaThrGluProArgTyrProProHisSerLeuSerTyrProVal 28
              |||:::||||||   |||      |||            |||            |||
Db       9254 GCCCGTACGTGGCCAGCAGCGCACCCGCCTCGATCGGATCACCCAGCGCCGTCCCCGTAC 9195

Qy         29 GlnIleAlaArgThrHisThrAspValGlyLeuLeuGluTyrGlnHisHisSerArgAsp 48
              :::            |||
Db       9194 CGTGCGCCTCCACCACAT------------------------------------------ 9177

Qy         49 TyrAlaSerHisLeuSerProGlySerIleIleGlnProGlnArgArgArgProSerLeu 68
                                                  |||:::||||||   |||:::
Db       9176 ------------------------------------CCACGTCGGCGGGCACCAGCCCCG 9153

Qy         69 LeuSerGluPheGlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSer 88
                                |||:::      :::         |||      |||   :::
Db       9152 CATCAGCCAGCGCCTCACGGATCACCCGCTGCTGCGCCGGACCACTCGGCGCCGTCAGCC 9093

Qy         89 HisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIleGluSerLysArgPro 108
              |||                                             |||   ||||||
Db       9092 CATTGC------------------------------------------TCGCACCGTCCT 9075

Qy        109 ArgLeuGluLeuLeuProAsp---------------------ProLeuLeuArgProSer 121
                 |||      ||||||                        ||||||   |||   |||
Db       9074 GA-TTGACCGCACTGCCCCGCACCACCGCCAGCACCCGATGACCGTTGCGGCGCGCGTCG 9016

Qy        122 ProLeuLeuAlaThrGlyGln---ProAlaGlySerGluAspLeuThrLysAspArgSer 140
                       :::|||         |||      |||                  |||:::
Db       9015 GAAAGCCGCTCCACCAGCAGCACACCCACACCCTCCGCCCACGCCGTCCCGTC-CGCACC 8957

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProPro------------------ 154
              |||      :::|||   |||               ||||||
Db       8956 CTCCGCGAACGCCTTGCACCGCCCGTCCGCCGCCAGCCCCCGCTGCCGCGAGAACTCCAC 8897

Qy        155 ---------------------HisThrAspPro------------GluLeuGluLeuVal 163
                                   |||      |||            ::::::   |||
Db       8896 GAACACCGTCGGCGTCGCCATCACCGTCACCCCACCGGCCAACGCCAGATCACACTCACC 8837

Qy        164 ProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGluIle 183
              ||||||            :::   |||   :::   |||   |||      |||
Db       8836 CCGCCG------------CAACGACTGCACCGCCAGATGCAACGCCACCAACGACGACGA 8789

Qy        184 ThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu---Glu 202
              |||                     ::::::   ::::::::::::   |||:::   :::
Db       8788 ACA------------------CGCCGTGTCCACCGTCAGCGCCGGCCCCTCCAACCCCAA 8747

Qy        203 GluAlaAlaLysPro-ProGlu---------ProGluLysProValSer------ProPr 217
                          ||| ||||||         |||      |||||||||      |||||
Db       8746 CGCATAGGCAACCCGACCGGAGATCACACTGCCGAGGGTCCCCGTGAGCACATGCCCGCC 8687

Qy        217 oProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLy 237
              |   :::   :::         |||   |||   ::::::         :::      ::
Db       8686 GTAGGTGTCGGCGGCGTCCCCCAGCTGGGTCGC-GTAGTCCTGCTGGGAGATGCCGCTGA 8628

Qy        237 s---------AlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLe 254
              :               ||||||   |||:::            ||||||         ||
Db       8627 AGACACCGGTCCGGGTGCCGCGCAGCGAGTCCGGAACGATGCCGGCCCGCTCCAGCGCCT 8568

Qy        254 uPro------LeuTyrAsnGlnPro---------------------SerAspThrArgGl 265
              ||||      |||      ||||||                              |||::
Db       8567 CCCACGACGTCTCCAGCAGCAACCGCTGCTGCGGGTCCATGGCCGTGGCCTCGCGCGGCG 8508

Qy        265 nTyrHisGluAsnIleLysIle------AsnGlnAlaMetArgLysLysLeuIleLeuTy 283
              :|||   |||   :::::::::         :::      ::::::            ||
Db       8507 ATATCCCGAAGAACTCCGCGTCGAACTCGGCCGCGGAACGCAGGAACCCGCCTTCCCGTA 8448

Qy        283 rPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAs 303
              |               |||   |||               :::      |||      ||
Db       8447 CATAGCTCTTGCCCGGCACCCCCGG------------CTCCGGGTCGTACAGCCCCTCGA 8400

Qy        303 pGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgAr 323
              |                     ::::::   ||||||:::
Db       8399 CGTCCCAGCCACGGTCGGCCGGGAACGGGGAGAGCGTGTGCCCGTCACCGGCCAC----- 8345

Qy        323 gAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLysGl 343
                                                   |||   |||::::::   :::::
Db       8344 -------------------------------------CAGGTCCCACAGGTCGTCCGGCG 8322

Qy        343 nArgGluLeuGlnGluArgMetGlnSerArg----------------------------- 353
              :|||            |||:::   :::|||
Db       8321 ACGCCAC---ACCGCCCGGGTAACGGCACGCCATGCCCACGATCGCGATCGGCTCACGCT 8265

Qy        354 -ValGly------GlnArgGlySerGly---------------LeuSerMetSerAlaAl 366
               ||||||      :::||||||   |||               |||::::::
Db       8264 CGTCGGCCATGACCGGCGTGGGGGCGGCCGCTGCCGGCTCCGCCTCGCTGTCGGGACCGA 8205

Qy        366 aArg------------------SerGluHisGluValSerGluIleIleAspGlyLeuSe 380
               :::                     |||   |||         :::::::::
Db       8204 ACAGTTCGTCGTGCAGGTGGAGGCGAGCGAGAGGGGGGTCGGGTGGTCGAA-------- 8153

Qy        380 rGluGlnGluAsnLeuGluLysGlnMetArgGlnLeu----------------------- 392
                              |||:::::::::||||||:::
Db       8152 ----------------GAGCAGCGTGTTCGGCAGGTTCAGGCCGGTGGCAGCGGTCAGGC 8109
```

```
Qy     393 -------------------------AlaValIlePro------ProMetLeuTyrAspAl 402
                                       |||   |||      |||
Db    8108 GGTTACGCAGTTCGAGAACGGTGAGGGAGTCGAACCCGGCGGACCG------------GA 8061

Qy     402 aAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro------Me 420
               |||   |||:::      :::         |||:::      ||||||      ::
Db    8060 AGGCCAGCTCCGGGTCGACCGCGTCGGC---CGAGGCGTGGCGCAGGATCCCGGCGGCCT 8004

Qy     420 tLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPh 440
           ::::   :::|||                           ||||||            ||
Db    8003 CCGAGCGCACAAG------------------------CGTGAGCAGAGCACGACGGCGTT 7968

Qy     440 eArgGluLys-------------------------------------------------- 443
           ||||   :::
Db    7967 CCGCCGGCGAGGGGCCGACGGAGAGCCACGGCCGGCCCTCGCCCCCGTCCTGTCCGCTCC 7908

Qy     444 ----------PheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu-- 459
                        :::      |||:::   :::|||         :::   :::
Db    7907 CGCCCGGTGCGCCGTCGCCCGGCCGCGTCCCTGGGGTGAGAGCGGCACGGTCGTCCGCGG 7848

Qy     460 -------------ArgLysThr-ValAlaGluCysValLeuTyrTyrTyrLeuThrLysL 475
                        |||:::|||    ||||||
Db    7847 CATGGCCGGCACGCGCCGAACGGGCGGCGGAG---------------------------- 7816

Qy     475 ysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnG 495
                   :::            |||||||||         |||      |||      |
Db    7815 --ACGCCGTCGAAGAGCGACCAGGTACGGCGGGTGACCGAACGCCCGACGAAGGTCTCCC 7758

Qy     495 lnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln------------------GlnGlnGlnG 509
           ||                  ::::::                              :::
Db    7757 AGTCCACGTCGGCGACGAGCAGGGAGCTCTCGCCGGCCGCGTGCGTCCCTGCGGCAGAGG 7698

Qy     509 lnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluA 529
                |||   |||   :::|||                  :::         :::
Db    7697 TCCCGCCGGCGCCCACCACGTCGTCGAGGACGTCGAGCGCGCTCGGGGTCCAGGGCCG 7638

Qy     529 laGluLysGluGluGluLys---------------ProGlu------------------- 537
                            |||               ||||||
Db    7637 TGACGCCGTCCCGTTCGAAGGACTTCGCCGCCGTTCCGGAGGCCATTCCGGCACCGGCCC 7578

Qy     538 --------------ValGluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspT 553
                         :::|||      :::      |||   :::   :::
Db    7577 ACGGCCCCCAGGCGATCGAGGCGGCACGGGCTCCCCTGGCCCGGCGCCGCTGCGCCAGGG 7518

Qy     553 hrSerGly------------------------------------------------GluA 557
             |||
Db    7517 CGTCGAGCGCGGCGTTGGCGGCCGCGTAGACGGCCTGCCCTCCGTTGCCCCACACGCCGA 7458

Qy     557 spAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyA 577
                :::||||||         ||||||   ||||||         :::|||
Db    7457 CGATGGAGGAGAAGAGCACGAAGGCGTCGAGCGGAAGGTCCGCGGTCAGCTCGTCCAGGT 7398

Qy     577 rgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleT 597
             |||                                                |||:::|
Db    7397 GACGG---------------------------------------------GCGGCGGTGA 7383

Qy     597 hrProGlnGlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrG 617
           ||      :::|||                  |||::::::|||||||||      |||
Db    7382 CCTTGTCCCGCAGTACGGCGTCCACCCGGTCGGAGGTGAGCGAGTCGAGC------ACAC 7329

Qy     617 luGluGluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaI 637
                            |||                     ||||||:::
Db    7328 CGTCGTCGAGCACACCTGCCGCGTGCAGTACGGCGACGAGGGGCCGCTCCTCGGGGAGGC 7269

Qy     637 leAlaArgMetValGlySer---------LysThrValSerGlnCysLysAsnPheTyrP 654
             ||||||         |||                  ||||||            :::|
Db    7268 CCGCGAGGAGCGCGGCCAGTGCGTCCCGGTCGGCGACGTCGCAGGCCTCGACGGC-CACT 7210

Qy     654 heAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGln---------------- 668
           ||      ::::::||||||   |||||||||
Db    7209 TCGGAACCGAGCGCCGTCAGCTCCTCGACGAGTTCGCCCACTCCGGGGCCGGCCGCGCCG 7150

Qy     669 --------HisLysLeuLysMetGluLysGluArgAsnAlaArg-ArgLysLysLysLys 685
                   |||:::         |||:::   |||   |||||| |||         :::
Db    7149 CGCCGGCCCACCAG---------GAGCAGACGCGCGGCGCCCGACCGGGCGAGCCTGCGG 7099

Qy     686 AlaProAlaAlaAlaSerGluGluAlaAlaPheProProValValGlu------AspGlu 703
           |||      |||            |||      |||||||||:::            |||
Db    7098 GCGACCTGCGCA------CCCAGGGCACCGGTGCCCCCGGTGATCAGTACGGTGCCGGAG 7045

Qy     704 GluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGluAlaGluAla 723
           :::      :::||||||   :::   :::|||                  :::
Db    7044 GACCACCGGTCGTCCGGCACGACCACGTCCGAGGTGTCGTTCCGTGCGTCCTCGCCGACG 6985

Qy     724 LeuHisAlaSerGlyAsnGluValProArgGlyGluCys---SerGlyProAlaThrVal 742
                 |||                  |||   |||      :::|||
Db    6984 GCGGTGGCGGGGGCCGTGGTGACGGTGCGAGCGGAGGCCGGGACGGGGACGGCGCGGG- 6926

Qy     743 AsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLysAspThr 762
           |||:::::::::|||   ::::::         |||:::   |||
Db    6925 AACCACACGGCGGACATGGACGCC---GGAGGGCCGTACCGCGAACTGGTTCTCTCCCCC 6869

Qy     763 GlyGlnAsnGlyProLys---------------ProProAlaThrLeuGlyAlaAspGly 777
           |||:::      |||:::               |||   ||||||            |||
Db    6868 GGCGAGCA-GCGCCCCGACGAACGTACGGAGCACCGCGGGCGACGCACTCCCGGGCAGGT 6810

Qy     778 ProProProGlyProProThrProPro---------ArgArgThrSerArgAlaProIle 794
           ||||||      |||         |||         ||||||      |||   |||
Db    6809 CCGCCAAGCCGCCCCAGCGCCGCCCCAGTTCGAGCGCGGCGACCTGTCCGAGTCCCCACA 6750

Qy     795 GluProThrProAlaSerGluAlaThr---GlyAlaProThrProProPro--------- 810
                    |||      :::   :::      |||      |||||||||
Db    6749 GCTGAGCGCCCCCGACGGAAGGGGTCTCGTCCGGCCGCACGGCCGCCGCCTCACGGTGG 6690

Qy     811 ---------------------AlaProProSerProSer--------------------- 816
                                   ||||||   |||:::
Db    6689 TGATCCAGAGCGGGGCTTCGGGTACCGCCTCGCCCAGCGCCTGGACGAGGGCGACGGCGT 6630

Qy     817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
              |||   |||      |||            |||               |||   |||
Db    6629 CGGCCGCGTCCGGCTCGGCCGGGGCGAGCCACAGAACT-CGGGTGCTGCGGCGGCACCG 6571

Qy     837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
                             :::   |||   |||:::            |||   ||||||
Db    6570 TCACCGGCGGCCCGCAACAGGTCCCCGGCTGCCTCCCGG------CCTGTACCGACCGGC 6517

Qy     857 LysAlaGluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLys 876
              |||         |||   :::               |||
Db    6516 ACGGCG------ACGGTGTGGACCTCGGCCCCGCGCACGGCGAGTTCACGTGCGATCGCG 6463

Qy     877 GlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLys 896
           |||         |||:::|||      ::::::   ||||||      :::
Db    6462 GGGGCGAGTCCGGAGTCGGCC---GTGTCCTCGTCCGGGGCCAGAAGCAGCACATGACCG 6406

Qy     897 GluGly---GlySerGlyArgAlaThrThrAla--------------------------- 906
              |||   |||      |||:::|||||||||
Db    6405 GCAGGCCCGGGGTTGCCGCGGTCGACGACGGCCGGCTGCCAAGCGACGCGGTAGCGGATC 6346

Qy     907 ------------------------------------LysSerSerGlyAlaProGlnAsp 914
                                                  |||:::|||   |||
Db    6345 CCGGCCGTCGTCGACCGTCCCCGCTCAGTGGCCCGAGCGTCCGCCGGCGAGCCGGCGGGC 6286

Qy     915 Ser--------------------------------------------------------- 915
           |||
Db    6285 TCGACGCCGGTCACGGCGGTGCGAACGGGGTCCAGCCAGTAGCGGTCGCGTTGGAAGGGG 6226

Qy     916 ------------------------------------------AspSerSerAlaThrCys 921
                                                        :::   |||:::
Db    6225 TAGGTGGGGAGGTCGATGCGCCGGCCGGTGCTCCCGGTGTGGAGTGCCGTGGCGTCCAGG 6166

Qy     922 SerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArg 941
              |||      |||:::                     |||         |||   |||
Db    6165 CCGGCGTCCCGGGTGAACACCGTCGCCAGCGCCGCCTCGAACACCTCCCGCTCCGCACGG 6106

Qy     942 ProSerLeuLeu---------ThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
```

```
              |||      :::         |||   |||      |||            ||||||
Db       6105 CCCCGCCGCATCGCCGGCACCACCACCACATCATCACCGGCCCCCAGGCACTCACCCGC- 6047

Qy        959 LysPro------------------LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAla 972
                 |||                        |||   |||
Db       6046 CATCCCCGTCAGCACCCCATGCGGACCCACCTCCACCAACGTCCCCACACCCAGACCACG 5987

Qy        973 IleProProIleGlnValThrLysValHisGluProProArgGluAspAla------Ala 990
                    |||      :::   :::   |||   ||||||      ||||||
Db       5986 CACCACCCCCACCCCATCCGCGAAACGCACCG-CCTCCCGCGCATGACGCACCCAATACC 5928

Qy        991 ProThrLysProAlaProProAlaPro------------------------ProProPro 1002
              |||      |||   ||||||   |||                        |||||||||
Db       5927 CCGGATCCCCCAACTCCCCCGAACCCACCACCCCACCCGACACACCCGACACCACCACCA 5868

Qy       1003 GlnAsnLeuGlnProGlu------SerAspAlaProGlnGlnProGlySerSerProArg 1020
                          |||:::               ||||||   ||||||:::   |||
Db       5867 CACCCGGCCGCACCCGACCGAACTCCAACGACTCCACAACCCCCCGGAACTCCCCCA--- 5811

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
                          :::|||||||||            |||      |||
Db       5810 ------------ACACCGGCTCCATCAACACCGAATGAAACCCATGCG------ACACAT 5769

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
                       |||   ||||||         |||            |||   |||
Db       5768 CCAACCGCCGGCACTCCACCCCATCCCCCATCAGCGAGGCAACCACCGACTCCA------ 5715

Qy       1061 GluValIleLysAlaSerProHisAlaPro---AspProSerAlaPheSerTyrAlaPro 1079
                          |||   ||||||   |||      |||         :::::::::|||
Db       5714 ------------GCACACCCACATCACCCGACAACACCACCGACCGCGGCCCATTCACCG 5667

Qy       1080 Pro-Gly---HisProLeuProLeuGlyLeu---------------HisAspThrAlaAr 1093
              |||       ||||||||||||      |||               |||         ||
Db       5666 CCGCAACCGACACCCACTCCCCCAACCCCTCAACAACCCCCCGCACCACCGACTCCGACG 5607

Qy       1093 gProValLeuProArgProProThrIleSerAsnPro----------------------- 1105
              ||||      |||   |||||||||         |||
Db       5606 CCCCCACCGACCACATCCCCCACCCACCGGCAACCCACCCATCAACCCACCCCGCGCCA 5547

Qy       1106 -------ProProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGl 1123
                     |||      |||:::      :::||||||      :::
Db       5546 CCACCAACCGCACCGCATCACCCAACGACAACACCCCCGCCACATACGCAGCAGC----- 5492

Qy       1123 yAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAlaLysAl 1143
                                             |||:::|||:::
Db       5491 -------------------------------CACCTCCCCCACCG-AATGACCCAACACC 5464

Qy       1143 aProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLysLeuAlaPr 1163
               |||      |||               |||   |||      |||            ||
Db       5463 ACCGACACCTCCA------------CACCCCGAGCCTCCAACGCCCGGAACAACGCCACC 5416

Qy       1163 oPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro-------- 1180
              |   :::         |||   :::   :::|||               |||
Db       5415 TCCAACGCAAACAGCGCAGGCTGAGCAAACTCCGTCCGACCCAAC-AACCCACCCACAGA 5357

Qy       1181 -------ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyTh 1198
                     ||||||            |||            ||||||
Db       5356 ACCGGAACCAGAACCGACCCCAGCACCCGCACCGGCACCCGCATCAGCCCCGGCACCCGC 5297

Qy       1198 rAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPr 1218
               :::      :::   |||      :::               |||||||||
Db       5296 GTCCACGTCGACGTCGCCGAACATCACATCCCGCAACGACCGACCATCCACCTCCCCCAC 5237

Qy       1218 oSerAspSer------------------------------------AlaIleThrTyrAr 1226
                  |||:::                                    |||         ||
Db       5236 CATCGACAACACCTCATCCAGCACCTCCGCAAACACCCCACCCCCCGCATACAACCCACG 5177

Qy       1226 gGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIl 1246
              |         ||||||      |||      |||            |||   |||
Db       5176 CCCCATACCAACCCACTGCGCACCCTGACCCGTG------AACAACACCCCCACC----- 5128

Qy       1246 eIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGl 1266
                                 :::|||   ||||||         |||:::||||||
```

```
Db       5127 -------CCACCCGGCACAGCCCGCCCGGACCGCACACCCCCAGACACCCTCCCGGCAGC 5075

Qy       1266 yHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerVa 1286
               :::
Db       5074 AAAC-------------------------------------------------------- 5071

Qy       1286 lThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAl 1306
                                          ||||||      |||||||||   |||   ||
Db       5070 ----------CCACCAAGAGCAGCGACGAGCTCCCCACGCCCACCACACGGCACCACCGC 5021

Qy       1306 aProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSe 1326
              |                              |||      :::   :::         ::
Db       5020 A------------------------------CGGTGCTCGAAAAGAGTCCGCGTCGAAGC 4991

Qy       1326 rIleGluGlyLeuMetGlyArgAlaIleProPro-------------------------- 1337
              :   |||            |||         |||
Db       4990 CAGCGAATAAGCCACATCCCGCGGATCCGCACCCACCCGCTCCACGTACTCACCCAACCG 4931

Qy       1338 ----------GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySe 1354
                           |||::::::|||            :::   :::         |||::
Db       4930 CACCGCCTGCGCCCGCAACGCCCCCTCCGACCGCGCGGACAACAACAGCGGCGACGGAGC 4871

Qy       1354 rIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGl 1374
              :            :::|||||||||                  |||   |||
Db       4870 AGCCGCGTCCGCGGTACCGCGGTCAGGAGCAGCGACGCCGTC-GATGTCCTCGACCCACT 4812

Qy       1374 uAlaLysLeuLeu------------------------------------LysArgGluGl 1382
                  :::||||||                                    :::|||
Db       4811 CCGGCGCCTCCTCGATGATCAGATGCGCGTTCGTCCCACTGATCCCGAACGACGACACAG 4752

Qy       1382 yThrPro---------------ProProProProProSerArg---------------As 1392
                  |||               ||||||   ||||||:::|||               ||
Db       4751 CCGCCCGGCGCGGGCGGTCCGACCGCCGGGGCCACCGGCGCGCCTCTGCCAGCAGGGCGA 4692

Qy       1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
              |                              ||||||:::::::::|||         ::
Db       4691 CGGTGCCCGACGACCAGTCGACGTGCGTGCTGGGCCGGTCCGCGTGAAGTGTCCGCGGCA 4632

Qy       1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
              :      |||         |||                                 :::
Db       4631 GCGTGCCGTGCTCCAGGGCAAG---------------------------------CAGCA 4605

Qy       1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
               |||   ||||||                  ||||||            |||   :::
Db       4604 TCTTGACCACACC------------CGCGGCCCCGC-------------GGCGGCCTG-- 4571

Qy       1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
                                          |||   :::      |||   ::::::   ::
Db       4570 ----------------------------GGTGTGACCGACGTTGGACTTCAGCGACCCGA 4539

Qy       1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
              :   :::   :::                                 |||||||||
Db       4538 GCCACAAGGGGTC------------------------------GCCCACCCGCTC----- 4514

Qy       1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
                           |||      :::||||||      :::         :::   |||||
Db       4513 ----ACGCCCGTACGTGGCCAGCAGCGCGCTCGCCTCGATCGGATCGCCCAGCGCCGTCC 4458

Qy       1512 oGlyThrAla------------------------------------------------Se 1516
              |   |||
Db       4457 CCGTACCGTGCGCCTCCACCGCGTCCACGTCACCCGGCGCGAGGCCGGCGTCGGCCAGGG 4398

Qy       1516 rSerSerGlyGlySerIleAlaArgGlyAla---------------------------Pr 1527
               ||||||      ::::::   |||||||||
Db       4397 CAGCTCGGATGACGCCCTCCTGCGCGGAGCGGTTGGGAGCGGTCAGGCCGTTGCTGGCAC 4338

Qy       1527 oValIleVal--ProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHis 1546
               |||:::|||  |||            ||||||   :::|||
Db       4337 CGTCCTGGTTGACCGCACTGCCCCGCACCACGCCCAGCACCC------------------ 4296

Qy       1547 GlyAlaProPheAlaGlyHisLeuProArgGlySerProValThrMetArg--------- 1563
              |||   |||   ||||||         |||   :::||||||   :::|||
Db       4295 GGTGTCCCGCACGCCGGGCGT---CGGAGAGCCGCTCCAGTGCCAGTACGGCGACACCCT 4239
```

```
Qy        1564 ------GluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGln 1581
                        |||      |||   :::      :::      |||   :::      :::
Db        4238 CGGACCAGGCCGTACCGTCGGCATCGGCGGAGAACGGCATGCAGCGGCCGTCGGGGGCGA 4179

Qy        1582 AspArgLysLeuThrSerThrProArg------GluIleAlaLysSerProHisSerThr 1599
                  |||            |||||||||                  ||||||
Db        4178 GCCCGCGCTGCCGCGAGAACTCCACGAACATGCCGGGCCCCGACATCACCGTGGCGCCGC 4119

Qy        1600 ValProGluHisHisProHisProIleSerPro--------------------------- 1610
                  ||||||      ||||||      ||||||
Db        4118 CCGCCAGAGCC-AGCCCGCAC------TCGCCCCGGCGCAGCGCCTGCACGGCCAGGTGC 4066

Qy        1611 ---------------TyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArg 1625
                                 ||||||   :::      ||||||
Db        4065 AGCGCCACCAGGGACGAGGAGCACGCGGTGTCGACGGTGAGC------------------ 4024

Qy        1626 SerHisIleProLeuAlaPheAspProThrSerIleProArg------GlyIleProLeu 1643
                              |||      |||      :::||||||         |||
Db        4023 ---------------GCCGGTCCGCCG------GTGCCGAGGGTGTAGGCGATCCGGCCG 3985

Qy        1644 AspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHis 1663
               :::|||      |||                  :::   |||         ||||||
Db        3984 GAGGCCACGCTCGCG------------------GTGGTGCCGGTGAGCAGGTACCCCTCG 3943

Qy        1664 LeuTyrPro------------------------ProTyrLeuIleArgGlyTyrPro--- 1674
                  ||||||                        ||||||               |||
Db        3942 ---TATCCGTCCGCTCCCTCGTGAAGCCTCGGCCCGTACTCCGGTGCGGTGGCGCCGACG 3886

Qy        1675 AspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGln 1694
               :::|||   |||         |||            :::         :::|||:::
Db        3885 AACACTCCGGCGTCGCCGCCGCGCAGTGACTCGGGGAGCACTCCGGCGTCCTCGAGGGCC 3826

Qy        1695 MetHisHisAsnThrAlaThr---------------------AlaMetAlaGlnArgAla 1707
                  |||      :::::::::                     |||:::|||   |||
Db        3825 TCCCACGACGCTTCCAGCAGCAAGCGCTGCTGCGGGTCCATGGCCGTGGCCTCGCGCGGC 3766

Qy        1708 AspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGly 1727
               |||:::               |||:::   |||      |||   |||   ||||||
Db        3765 GATATC---------------CCGAAGAACTCC------GCGTCGAACTCGGCCGCC--- 3730

Qy        1728 ProArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThr 1747
                              :::            |||            ::::::|||
Db        3729 ---------------GAGTGCAGGAACCCGCCCTCCCGCACATAGGTCTTGCCGGGCGTC 3685

Qy        1748 ProGly------------ThrProAlaThrAlaMetAspArgLeuAla---------Tyr 1760
               ||||||            :::|||:::   :::      |||
Db        3684 CCGGGCTCGGGGTCGTACAGCCCGTCGAGGTCCCAGCCGCGGTCGGTGGGGAAGGACGAG 3625

Qy        1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
               :::   :::||||||         ||||||:::   ::::::         :::|||
Db        3624 ATGGCGTCGGCGCCGGACTCCAGCAGCTCCCAGAGAGCGGCGCGCGAGTCGACGCCGCCG 3565

Qy        1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
               |||            |||   ||||||            |||         |||
Db        3564 GGGAGCCGGCAGCCGACA---CCCACGATGGCGATCGGGTCGTCGTCGTCCCGGCTGCCG 3508

Qy        1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
                  |||      |||               ::::::|||                  |||
Db        3507 CCCCGGGCGTCCGACACCCCTACCGCGGCCGACGAGAGCACTCCGGGCCCACCGGCAACC 3448

Qy        1821 ValGluHisAlaPro------------IleTrpArgProGlyThrGluGlnSerSerGly 1836
                           |||                           |||   :::      |||
Db        3447 GAGTCCTTCGGCCCGCCCGCGGCTTCCACGATGTGCTCGGCGACGGCACGCGGGGTGGGC 3388

Qy        1837 SerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHis 1856
                  |||   :::|||            ||||||         |||   |||   ||||||
Db        3387 AGGTCGTAGACCAGCGTGGCGGCGAGGGGGAGCCCCGAGGCGCCCTTCAGCAGGTCGCAC 3328

Qy        1857 -AlaHisGlnHisSerProIleSerPro------------------------ArgThrGl 1868
                |||         ||||||   ||||||                        |||   ::
Db        3327 AGCTCGACCGCGGTCACCG---AGTCCATGCCCTGGGAGGTGAAGGTGCGGTCGAGCGCG 3271

Qy        1868 nAspAlaLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleTh 1888
               :            :::|||
Db        3270 ACCTCGCCGGGATCGTCGT----------------------------------------- 3252

Qy        1888 rAlaValGluProSerLysProThrValLeuArgSer------------------ThrSe 1902
                         |||   :::|||         |||
Db        3251 -------GGCCCAGGACGGCCGCGGTGCTCTCGCGGACCAGGCGGACGGCATCCTCGGGG 3199

Qy        1902 rThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGl 1922
                ||||||:::||||||   |||:::         ||||||         ||||||   ||
Db        3198 GACATCGGCTCCGGTGACGCCGAGT---------CCGCCTCGTCGTCCGTGTCCCTGTGG 3148

Qy        1922 yGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLysGluAl 1942
               |   :::                                          |||:::   ::
Db        3147 GACGTCC------------------------------------------CCTCGGTGCTC 3130

Qy        1942 aProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLy 1962
               :|||         |||         ||||||            |||   :::   :::
Db        3129 TCCGGGTCCGGGGAGACGCTGTGCCCCACGG------------CATCGGTGGCGGTCACA 3082

Qy        1962 sProProAlaArg---------------------------------SerGlyLeu----- 1969
                ||||||   |||                                 ||||||
Db        3081 CCCGCCCCGGCGTCGGCCGTGACCGGCCGGCTCAGCGCGGGCGACCAGTGGGTACGGCGT 3022

Qy        1969 ------------------------------------------------------------ 1969

Db        3021 TGGAAGGGGTAGGTGGGGAGGTCGATGCGCCGGCCGGTGCTCCCGGTGTGGAGTGCCGTG 2962

Qy        1970 -GluProAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSe 1989
                   |||         |||               |||   |||         |||   ::
Db        2961 GCGTCCAGGCCGGCGTCCCGGGTGAACACCGTCGCCAGCGCCGCCTCGAACACCTCCCGC 2902

Qy        1989 rGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerPr 2009
               :   |||         |||   :::||||||            |||||||||      ||
Db        2901 TCCGCACGGCCCC---GCCGCATCGCCGGCACCA------CCACCACATCATCACCGGCC 2851

Qy        2009 o---------AspProProAlaProProAlaSerAlaSerAsp----------------- 2020
               |            ||||||:::|||   |||      :::|||
Db        2850 CCCAGGCACTCACCCGCCATCCCCGTCAGCACCCCATGCGGACCCACCTCCACCAACGTC 2791

Qy        2021 -ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSe 2040
                ||||||   :::               |||   |||   :::            |||::
Db        2790 CCCACACCCAGACCACGCACCACCCCCACCCCA-TCCGCGAAACGCACCGCCTCCCGCGC 2732

Qy        2040 r------LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerPr 2058
               :            |||   ||||||      ||||||         ||||||   :::||
Db        2731 ATGACGCACCCAATACCCCGGATCCCCCAACTCCCCC---------GAACCCACCACCCC 2681

Qy        2058 oValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAs 2078
               |      :::|||      |||         |||         |||      ||||||
Db        2680 ACCCGACACACCCGACACCACCACCACACCCGGCCG-------CACCCGACCGAACTCCA 2628

Qy        2078 pLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGl 2098
                :::   |||                                 |||   :::|||
Db        2627 ACGACTCCAC------------------------------AACCCCCCGGAACTCCCCCA 2598

Qy        2098 yGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerPr 2118
                         |||         ::::::|||:::         ||||||         ||
Db        2597 ACACCGGCTCCATCAACACCGAATGAAACCCATGCGACACATCCAACCG---------CC 2547

Qy        2118 oLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHi 2138
               |   |||         |||      :::|||:::   |||:::            |||||
Db        2546 GGCACTCCACCCCATCCCCCATCAGCGAGGCAACCACCGACTC------------CAGCA 2499

Qy        2138 sIleSerGluValIleThrGlnAspTyrThrArgHisHisPro------GlnGlnLeuSe 2156
               |            ||||||:::   :::   |||      |||      ::::::
Db        2498 CACCCAC------ATCACCCGACAACACCACCGACCGCGGCCCATTCACCGCCGCAACCG 2445

Qy        2156 rAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLe 2176
                   |||||||||   ||||||   :::   |||                        ||
Db        2444 ACACCCACTCCCCCAACCCCTCAACAACCCCCCGCACCACCGA---------------CT 2400

Qy        2176 uArgArgProProSerAspLeuTyrLeuPro---ProProAspHis-------------- 2190
```

```
               ||||||||||            :::   |||   ||||||   |||
Db        2399 CCGACGCCCC------------CACCGACCACATCCCCCACCCACCGGCAACCCACCCA 2352

Qy        2191 ----GlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluPr 2209
                         |||            ||||||            ::::::   |||
Db        2351 TCAACCCACCCCGCGCCACCACCAACCGCACCGCATCACCCAACGACAACACCCCCGCCA 2292

Qy        2209 oAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGl 2229
                   ::::::|||   |||                                    |||
Db        2291 CATACGCAGCAGCCACCTC------------------------------------CCCCA 2268

Qy        2229 uGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeu----------- 2245
                   ||||||:::      |||   |||            |||   |||
Db        2267 CCGAATGACCCAA------CACCACCGACACCTCCACACCCCGAGCCTCCAACGCCCGGA 2214

Qy        2245 ------------------------------------------------------------ 2245

Db        2213 ACAACGCCACCTCCAACGCAAACAACGCAGGCTGAGCAAACTCCGTCCGACCCAACAACC 2154

Qy        2246 ----TyrArgAspGlyGluGlnThrGluProSer---ArgMetGlySerLys-Ser---- 2261
                   :::|||   |||   :::|||:::||||||   |||   |||:::::: |||
Db        2153 CACCCACAGAACCGGAACCAGAACCGACCCCAGCACCCGCACCGGCACCCGCATCAGCCC 2094

Qy        2262 -----ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnS 2280
                    |||   :::||||||      |||               :::|||      |||
Db        2093 CGGCACCCGCGTCCACGTCGACGTCGCCG------------AACATCACATCCCGCAACG 2046

Qy        2280 erAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgA 2300
                             |||         :::                  :::   |||
Db        2045 ACCGACCATCCACCTCCCCCACCATCGACAACACCTCATCCAGCACCTCCGCAAACACCC 1986

Qy        2300 snGluProGluTyrAsnIleSerGlnProGlyThrGlu---------------IlePheA 2315
                    |||   ||||||               |||                  :::   |
Db        1985 CACCCCCCGCATACAACCCACGCCCCATACCAACCCACTGCGCACCCTGACCCGTGAACA 1926

Qy        2315 snMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluH 2335
               ||   |||         ||||||
Db        1925 ACACCCCCACCCCACCCGGCACAGCC---------------------------------- 1900

Qy        2335 isAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrA 2355

Db        1900 ------------------------------------------------------------ 1900

Qy        2355 spGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerA 2375
                 :::   :::   :::||||||   :::      |||   ||||||      |||:::
Db        1899 --CGCCCGGACCGCACACCCCCAGACACCCTCCCGGCAGCAAACCCACCAAGAGCAGCGA 1842

Qy        2375 laSerLeuPro------------------------------------------------- 2378
                 |||   |||
Db        1841 CGAGCTCCCCACGCCCACCACACGGCACCACCGCACGGTGCTCGAAAAGAGTCCGCGTCG 1782

Qy        2379 --AlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProG 2398
                 |||:::         |||:::                     :::      ||||||
Db        1781 AAGCCAGCGAATAAGCCACATCCCGCGGATCCGCGCCCACCCGCTCCACGTACTCACCCA 1722

Qy        2398 lyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProA 2418
                             |||:::            ||||||   |||   |||   :::   :
Db        1721 ACCGCACCGCCTGCGCCCGC------AACGCCCCCTCCGACCGC---GCGGACAACAACA 1671

Qy        2418 laProGlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyA 2438
               ::   |||   :::      :::               :::   |||      ||||||
Db        1670 GCGGCGGGACGTCCGGAAGCGAGGCGGCGACGGAGACGGCCGGGGTGGGCCGGGAGGGCA 1611

Qy        2438 spCysAsnArgArgThr---------------ProLeuThrAsnArgValTrpGluAspA 2453
                 |||:::|||   |||               |||:::   |||         |||
Db        1610 GGTGTTCCAGGACCACATGGCAGTTGGTGCCACCGATGCCGAAC---------GAACTGA 1560

Qy        2453 rgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnA 2473
                 |||:::|||   |||:::   |||
Db        1559 CACCCGCCAGCAGGGGTGCGCCGCCC---------------------------------- 1534

Qy        2473 laGlyValMetAlaSerProProProPro-GlyLeuPro---AlaGlySerGlyProLeu 2491
                 |||         |||                   |||   ||||||:::||||||
```

```
Db        1533 --GGGGCGAGCGGCAGCTCGGCGGCAGCGGTCTGCACCTTCAGCCGGAGCTGGTCCAGCG 1476

Qy        2492 AlaGlyProHisHisAlaTrp 2498
                     |||         |||
Db        1475 GGATGGCCGGAGAGGGGGTGG 1455

RESULT 12
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.:              8.85e-19      Length:        4403765
Score:                  668.50        Matches:       617
Percent Similarity:     32.38%        Conservative:  255
Best Local Similarity:  22.91%        Mismatches:    1044
Query Match:            5.06%         Indels:        793
DB:                     3             Gaps:          113

US-09-522-753-5 (1-2517) x US-09-103-840A-2 (1-4403765)

Qy           4 SerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrProProHisSer 23
               ||||||:::               ||||||   |||      |||   ||||||   |||
Db     2794906 TCGACGGAAGCCCGCATCGTCTGCTGGCGGCGCACCGCAATACGCGATCCACCGGTCAGC 2794965

Qy          24 ---------------------LeuSerTyrProValGlnIleAlaArgThrHisThrAsp 36
                                    ||||||               ||||||      |||
Db     2794966 AGGCGGAACCGGTCATGCAAGCTGTCGACGATTTCGGTCAGCGCCAGGGCACGCACC--- 2795022

Qy          37 ValGlyLeuLeuGluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGly 56
                                             ||||||   ||||||   ::::::   |||
Db     2795023 ------------------------------CGCGACGCTGCAAGTTCGATCGCCAGCGGA 2795052

Qy          57 SerIleIleGlnProGlnArgArg-ArgProSerLeuLeuSerGluPheGlnProGlyAs 76
                           |||      |||       |||                  ||||||
Db     2795053 ---------ATGCCGTCGAGTCGGTGGCAGATCTCGG------------TCACCAGGGCG 2795091

Qy          76 nGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGl 96
                            :::|||         |||         |||   :::   |||
Db     2795092 AGTTGTCGGCAGTGATCTCGAGTTCGGGCCGCGCCTCACGAGCGCGGTCGGTGAA----- 2795146

Qy          96 yLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspPr 116
                                                    :::|||:::   |||      ||
Db     2795147 -------------------------------------CAACTCGATCGCCTCGCCGTGCC 2795169

Qy         116 oLeuLeuArgProSerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuTh 136
               |               |||   |||   ||||||
Db     2795170 CCAGCGGGGGAACCCGCCAAATCTGCTCACCGGCCAC----------------------- 2795206

Qy         136 rLysAspArgSerLeuThrGlyLysLeuGluProValSerProProSerProProHisTh 156
                :::||||||         |||:::
Db     2795207 -CGCGATCGGTTCCCGGCTGGTCGC----------------------------------- 2795230
```

```
Qy        156 rAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMe 176
                                    |||   |||   ::::::   |||
Db    2795231 ----------------CAATACCCTCAGCGCTGGGCACGCCCCGAG-------------- 2795260

Qy        176 tAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLy 196
                  |||   |||:::            |||||||||   ::::::   ::::::
Db    2795261 -CAACGCGACGATCAGAGCCGCGCACCCGTCGAGCAAGTGCTCGCAGTTGTCCAGCACTA 2795319

Qy        196 sGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSerPr 216
               |||                                             ||||||:::
Db    2795320 CCAG---------------------------------------CATGCGCCGGTCGCCGA 2795340

Qy        216 oProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLy 236
                  |||            ||||||         ::::::|||
Db    2795341 TACGCCGCACAATGGTGTCCACCGTCGAGCGGCCCGGCTGATCCGGCAACCCCAAAACCC 2795400

Qy        236 sLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuPro-- 255
               :::            |||:::         |||         ||||||::::::|||
Db    2795401 GCGCCGCCGCGATCGGCACCAGCGCCGGGTCGGTGATCGGCGCCAGGTTGACATACCAAA 2795460

Qy        256 ----------------------------------------------------LeuTyrAs 258
                                                                        ::
Db    2795461 CCCCGTCCGGATAACCGTCGGCAACGGCGCTCGCGACCTGTGTCGCCAGGCGTGTCTTTC 2795520

Qy        258 nGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLy 278
              :   |||:::||||||   :::                  :::   |||
Db    2795521 CGACCCCGCCGACACCGGTAAGGGTGACCCACCGTTTGACGTCCAGCAGCCCACGGAC-- 2795578

Qy        278 sLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPh 298
                     :::   :::|||            |||         :::
Db    2795579 -------TTGCGCCACTTCGTCGACGCGCCCCACCAGCCGAGTGAG-------------- 2795617

Qy        298 eCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAs 318
                                             |||   :::|||      |||      ::
Db    2795618 -------------------------------CTGGGCCGGAAGACAGTGCGCACCAACGA 2795646

Qy        318 nAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPhePr 338
              :   |||      :::   :::|||      |||:::            :::
Db    2795647 CTTTCCGGGTCCGCAGCGGCGGGAACGCGTTGTGCAGATCAGGGTGACACAGCTGCACCA 2795706

Qy        338 oGluIleArgLysGlnArgGluLeuGln------------GluArgMetGlnSerArgVa 354
                     |||      :::   :::|||            |||   :::|||   |||||
Db    2795707 CCCGTTCCGGTCGGGGCAGGTCGTCCAGCCGGTAGGTACCGAGGTCGTTCAGCCACGCGT 2795766

Qy        354 l---GlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSe 373
              |   ||||||                                          :::|||||
Db    2795767 CCTTGGGCAG------------------------------------------CAGGTCAG 2795784

Qy        373 rGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAl 393
               |     |||   |||   |||   |||      |||         :::   ||||||||
Db    2795785·CAACCAGATCGCTGGTAGTTCCCGACAACACGGTCTG---GCCCCCGTGGCCAGCTCGC 2795841

Qy        393 aValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIle-------- 410
              |      |||            |||               :::      :::
Db    2795842 GCAG---CCGGGCGGTGCGGTCGATGGTCGGCCCTACGCAGTTGCCCTCGTCGGGTGACG 2795898

Qy        411 -AsnMetAsnGlyLeuMetAlaAsp---------------------------ProMetLy 421
               ::::::   |||:::   ||||||                           |||:::::
Db    2795899 ACACCTCCCCGGTGTGCATGCCGATGCGCAGCCGGATCGGTGCCAGCGGCGCCCGCTGCA 2795958

Qy        421 sValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheAr 441
              :            |||      :::            |||:::   :::|||
Db    2795959 AGCCCAGGGCGCACGCCACGGCGTCGGATGCGCGGGCGAACGCCACCAAGAAGCTGTCGC 2796018

Qy        441 gGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLy 461
                                                |||         |||   :::|||
Db    2796019 CTTCGCCCTGTTCGACCGGGCAAACCCCGCGGTGCTCGCGAACCAATTCGGTCAGCGTTC 2796078

Qy        461 sThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSe 481

Db    2796079 G----------------------------------------------------------- 2796079
```

```
Qy        481 rLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGl 501
                  |||:::         |||||||||   |||   :::   :::|||      :::
Db    2796080 ----GTCCAGTTTGGCGATCGCCGTCGTGTCAAGCTGAGACCCCGGCAGGTGGGTCGCGC 2796135

Qy        501 nGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAs 521
                           |||||||||:::   |||   |||
Db    2796136 CCTCGATATCGGCCAGCAGCAACGTCA-CCGTGCCCG----------------------- 2796171

Qy        521 pGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsnAs 541
                           :::      :::                  |||:::|||
Db    2796172 -------TCGGTACAAGCTCGCTCACACCATCTGCGCTCCAGTCCACAGGTACCACGT-- 2796222

Qy        541 pLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLy 561

Db    2796222 ------------------------------------------------------------ 2796222

Qy        561 sGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyAr 581
                                                            ||||||:::|||
Db    2796223 ----------------------------------------------CGACGCCGGGGTGA 2796236

Qy        581 gIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGlnSe 601
               |||         :::|||                     :::|||   |||   :::||
Db    2796237 -ATCTTGCTCATGCTAGCC---------------------AGCATCGAGCCAGCGCGTAG 2796274

Qy        601 rAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet-- 620
              ||||            :::   |||:::   |||                     |||
Db    2796275 CGCATTACATCGGCACCTGCGCCTAGATTGCTCGAAATCTCTTGGCCGCCGGTCCATGTG 2796334

Qy        621 ----------------GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSe 635
                                       ::::::               |||      :::||
Db    2796335 TTCTACGCGCTTTAGTCGATGCATTCGGCGACCGGCGTGCCATCGCGGCGG-ACCTACAG 2796393

Qy        635 rAlaIleAlaArgMetValGlySerLysThr---ValSerGlnCysLysAsnPheTyrPh 654
              |      ||||||   |||      :::      |||   ||||||
Db    2796394 T------GCCCGTGCTGTCCGCTGGCAATTGTGAGTCCCCCAGTGC-------------- 2796433

Qy        654 eAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGl 674
                                                      ||||||:::   :::
Db    2796434 -------------------------------------TGGCAGCATCGCCCGCAAGAACC 2796456

Qy        674 uLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAl 694
                        :::      |||         |||:::||||||      |||   |||::
Db    2796457 GACACGACCGCATCGTGGGCGGTG-CCGTCGAAGTCGCCGGCTGACCGATCGGCGGAGTC 2796515

Qy        694 a------------AlaPheProProValValGluAspGluGluMetGluAlaSerGlyVa 710
              :               |||      :::      |||::::::      |||:::
Db    2796516 ACCGGCCCGATGGGGTTTCCGAAGGCTAGGGAATGATGACGATGGGGCGGCCGCCTCGGC 2796575

Qy        710 l------------------------SerGlyAsnGluGluGluMetValGluGluAlaGl 722
                                       :::   ||||||      :::
Db    2796576 CGCCTTCGCCGTAACCCCCAACCATGCGGAAAACGAGCCTAGCGTCGCCCGGCCGCGCAG 2796635

Qy        722 uAlaLeuHisAlaSerGly---AsnGluValProArg------------------GlyGl 735
               |||   |||      |||   |||:::      |||                  |||
Db    2796636 AGCGAGCCATCGCGGTGGCGCCAACGACAGGAAGCGATCCGGATTCTCTGACCATGGTGG 2796695

Qy        735 uCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHi 755
               |||||||||      |||:::         ::::::   |||      |||||||||
Db    2796696 GTGTTCTGGCTACGTGACGTTA---------ACGGAGATGGAGGGGCCGCCTTCGCCG-- 2796744

Qy        755 sThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAl 775
                                                |||   ||||||
Db    2796745 ----------------------------------CCTTCACCGCCG-------------- 2796756

Qy        775 aAspGlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleGl 795
               :::   |||   ||||||   |||   |||                  :::|||   ||
Db    2796757 -GAACCGCCGGAGCCAGGGTCGCCCCTCCCGTTGCCGGAGCCACCCGACTCGCCCGACGA 2796815

Qy        795 uProThrProAlaSerGluAlaThrGlyAlaProThrPro------ProProAlaProPr 813
              ||||||||||                  ||||||   |||      ||||||   |||||
Db    2796816 GCCGACGCCGCCGGAGGTCAAGCCACCGGCACCGCGTCCGCCGTCACCTCCGCGCCCGCC 2796875

Qy        813 oSerProSerAlaProProProValValProLysGluGluLysGluGluGluThrAlaAl 833
```

```
              |||||||   :::|||||||||::::::   :::            |||         ||
Db    2796876 GTCCCCGCCGTCACCGCCGCCGATGCTGCG-AGGCGGAGGGGCGCCGAAGCCGCCGGAGC 2796934

Qy        833 aAlaProPro------------ValGluGluGlyGluGluGlnLysProProAlaAlaGl 849
              ||||   |||            |||      |||               |||
Db    2796935 CGCCGGTCCCGCCGTCGCCTCCGTCACCACCGGGGGCGCCACCGTCTCGCCCGGCCCCAC 2796994

Qy        849 uGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGl 869
               :::   ||||||            |||      |||      :::      ::::::
Db    2796995 CCAAGCCGCCGTT---GCCGCCGTTGCCACCCGGCCCGCCGTCGCCTGCATCAGCAAAGC 2797051

Qy        869 uAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGl 889
               |||         |||                        |||      ||||||   ||
Db    2797052 TGCCGTTGCCGTCCCCACCGTGACCGCCGTTCCCGCCGTCGCCTCCGTCACCGCCGGGGG 2797111

Qy        889 yAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThr------------Th 905
              ||||               :::   |||   :::   |||||||||            ||
Db    2797112 CGCC---------------GAAGCCGGCCTTGCCGTGCGCGCCACTTGTGGAACCGAAAC 2797156

Qy        905 rAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGl 925
              ||||   |||:::||||||         :::   ::::::   |||      |||
Db    2797157 CGCCTTGTCCGCCGGGGCGGCCCCACCCGCCGTCGCCGCCGTCACCTCCGTCGCCGCCAG 2797216

Qy        925 uValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer-ProArgProSerLeuL 945
               :::         :::      :::      :::   :::::: |||:::|||   |||
Db    2797217 GCTC---TCCGTCAAAATCCGCGAGATAGGTAAAGCCGTCACCGCCCAAGCCACCATTAC 2797273

Qy        945 euThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysG 965
                            ||||||   :::      ||||||      |||         :::
Db    2797274 CAGCGTCCCCGCCCGACCCGCCGTCACCGCCGTCCCCGCCAACGCCT---------CGAT 2797324

Qy        965 lnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProP 985
                            |||   |||   ||||||                        ||||
Db    2797325 TGCCGACCTCGCCGGCGGGTGCCGACCCGCCG--------------------GCCCCGC 2797363

Qy        985 roArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsn- 1004
              ||         |||   |||      |||   ||||||:::||||||   |||   :::
Db    2797364 CGTTTCCGCCGGCGCCGCCCCACCCGCCGTAGCCACCGTCGCCGCCGTCGCCGCCGTCGC 2797423

Qy       1005 --------------------------------------LeuGlnProGluSerAspAlaP 1012
                                                    |||   |||         ||||
Db    2797424 GGCCCGTCGTTTCGTTAATGTCAAAGCCGTCAACGCCGTTACCGCCGACCCCACCAGCCC 2797483

Qy       1012 roGlnGlnProGlySerSerProArgGlyLysSerArg---------SerProAlaProP 1029
              ||   :::|||   :::   |||         |||               ||||||||||
Db    2797484 CGCCTAGGCCTCCGGCCCCGCCGTCACCACCGTCGCCGGTCTGAGTTCCGCCGGCGCCAC 2797543

Qy       1029 roAla---------------------------------------AspLysGluAlaPheA 1036
              |||||                                             :::   |||
Db    2797544 CGGCCCCGCCGTCGCCTCCCGCCCCACCGCTGCCGCCGTCGAAGCCGTCGAAGCCCTTTA 2797603

Qy       1036 laAlaGluAlaGlnLysLeuProGlyAspPro------------ProCysTrpThrSerG 1052
                :::|||:::         |||      |||            |||   |||
Db    2797604 GGTCGGAGTCGGGCGACCAACCCGCGCCACCGGCCGCGCCGTTGCCTCCCTGGCCG---- 2797659

Qy       1052 lyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspP 1072
                         ||||||||||||         |||   |||||||||   ||||||
Db    2797660 -----------CCGGTTCCGCCGCCGCCGTTCATCCCGGCGTCGCCGCCCGCCCCGCCG- 2797707

Qy       1072 roSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrA 1092
                                  ||||||      |||   |||   ||||||
Db    2797708 -----------------TGTCCACCAACCCCGCCGCCGCCGCCGGGGCTG---------- 2797740

Qy       1092 laArgProValLeuProArgProProThrIleSerAsnProProPro-LeuIleSerSer 1111
                   |||      |||   ||||||               |||||| |||      |||
Db    2797741 -----CCGCCCCGGCCAGCCCCACCTTGG------------CCGCCGGCTCCGCCGTTCC 2797783

Qy       1112 AlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerVal 1131
                 :::   |||   :::|||   |||:::
Db    2797784 CGCCGTCGCCCAGAAATGCTCCGCCGGCGC------------------------------ 2797813

Qy       1132 GlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGly 1151
                    |||   |||:::      |||         |||:::   |||   |||:::
Db    2797814 ------CACCAGCCCCACCGGCGCCACCA-GCCCCACCGTTGCCGCCAGCAACGGTGAGC 2797866

Qy       1152 LeuPro---LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGlu 1170
                 |||      |||      |||            |||   |||
Db    2797867 CCTCCGAGGGCACCGTGCGCGCCGTCGCCACCCTTGCCGCCGTCA--------------- 2797911

Qy       1171 GlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGln 1190
                    ||||||         |||   ||||||            :::|||   |||
Db    2797912 CCGCCGTCACCGATGTCGCCGGCGTCACCGCCC------------TTGCCTCCAGCCCCA 2797959

Qy       1191 GluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLys 1210
                 |||      :::         :::   ||||||         |||   ||||||
Db    2797960 CCGGCCCC-GCCATCACCGCCGAGAGCTTCGGCAGCGGTGCCGTCGGCCCCATCACCACC 2798018

Qy       1211 GlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThr 1230
              |||                  |||      :::|||:::|||         ::::::
Db    2798019 GGCTCCGCCGTCCCCGAATAGCCCGGCGTTGCCGCCGTCACCGCCCTGGCCGCCGTCGCC 2798078

Qy       1231 HisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAsp 1250
                 |||                        ||||||:::
Db    2798079 GCCGGCCGCGGCGGC------------CTTGGCACCGTTGCC------------------ 2798108

Qy       1251 SerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyr 1270
              :::   :::   :::      |||:::                     |||   :::
Db    2798109 GCCGACGCCGCCGTCGCCGCCGGTCAGTGGCCCGTGTTTGCTGGCGTCCACGCCGTT--- 2798165

Qy       1271 GluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSer 1290
                 |||:::   |||   |||   :::               :::   |||   |||:::
Db    2798166 ---GGCCGCGGAGGTGCCGTTGCCGCTGTCACCCCCAGACCGCCGCGACCGCCTGCGCC 2798222

Qy       1291 LysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThr 1310
                                                     ||||||         |||:::
Db    2798223 --------------------------------GGGGTCACCGCCGTTACCGCCCGCTCC 2798249

Qy       1311 TyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeu 1330
                          :::|||            |||:::   :::|||
Db    2798250 GCCGGCGCCGCCGACGGTGATACCAATGCCGCCGTT---GCCGCCGGCCCCGCCAACGCC 2798306

Qy       1331 MetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHis 1350
                 |||   |||            ||||||   ||||||      ::::::
Db    2798307 GCCGGCGCCGCCGA-GTCCGCCGTCGCCACCGACCCCACCGGTGCCGTGACTGCCGA--- 2798362

Qy       1351 IleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyr 1370
                                         |||         |||      :::      :::
Db    2798363 ---------------------------CCGTCCCCGAAGGTGCGGCCCCGCCGACCCCAC 2798395

Qy       1371 LeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSer 1390
                    |||               :::|||            |||||||||||||||:::
Db    2798396 CGTCCCCGCCATGTCCACCGACCCCGCCGGCACCG-CCATCGCCGCCACCACCGCCGGCC 2798454

Qy       1391 ArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAla 1410
                    :::                        |||   ||||||
Db    2798455 CCACCGGTGCCGCCGATACTGTCGATA---CCGTTGGCGCCCCTG--------------- 2798496

Qy       1411 HisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArg 1430
                                                                    |||
Db    2798497 ---------------------------------------------------GCCCCGGCC 2798505

Qy       1431 GluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGly--- 1449
                    |||      ||||||   ||||||   ||||||   ||||||      |||
Db    2798506 CCACCGCTAGCGCCCACACCGCCGTTGCCGCCGGCCCCG---CCGTTGCC-GCCGGCACC 2798561

Qy       1450 SerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLys 1469
              ::::::|||:::      |||            |||||||||:::
Db    2798562 GCCGTCACCCGACACCGACCC------------ACCGGCGCCACC--------------- 2798594

Qy       1470 LysHisAspValArgSerLeuIleGlySerProGlyArgThr-PheProProValHisPr 1489
                                      |||:::   |||   |||    ||||||      ||
Db    2798595 ------------------------GGCACCACCGGCACCACCGGCACCGCCGGCCTGCCC 2798630

Qy       1489 oLeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLy 1509
              |
Db    2798631 C----------------------------------------------------------- 2798631
```

```
Qy         1509 sSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIl 1529
                 :::   |||                  |||         ||||||   :::
Db      2798632 -GCGTCGCCCTGACCCCGTTGCCTCCCGGTTGGCCGAGGGCGAGGGCATCT-------- 2798682

Qy         1529 eValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaPr 1549
                       |||   |||   |||      :::||||||                     ||
Db      2798683 -------GAACCAGGCGCGCCCGAATTGGCCCCGTTGGCGCCGGCCGCGCCATCGCCACC 2798735

Qy         1549 oPheAlaGlyHisLeuProArgGlySerProValThrMetArg---------GluProTh 1566
                |            ||||||         |||         |||            |||::
Db      2798736 A------------TTGCCGCCGGCGCCACCATCGCCGACCCGGCCGGCATTGCCGCCGTC 2798783

Qy         1566 rProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuTh 1586
                :|||   |||
Db      2798784 GCCTCCGTTGCCC---------------------------------------------CC 2798798

Qy         1586 rSerThrProArgGluIleAlaLysSerProHisSerThrValProGluHisHisProHi 1606
                 :::   |||      :::|||   :::|||         :::|||         |||
Db      2798799 GGCGCCGCCCGCGACGCTGGCTTGCGCACCGTTGCCACCGTTACCACCGTTGCCGCCGCT 2798858

Qy         1606 sProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSe 1626
                 |||      |||:::                  ::::::         |||      ::
Db      2798859 GCCG---GGCCCGTGG---------------TCGCTGGCGTCCACACCGCTG---GCCGC 2798897

Qy         1626 rHisIleProLeuAlaPheAspProThrSerIlePro-ArgGlyIleProLeuAspAlaA 1646
                :   :::||||||         |||         ||| ||||||:::|||
Db      2798898 GCGGGTGCCGTTGCCGCTGTCGCCGCCCAAG---CCGCCGAGGCCTCCCGCGCCGGGGTC 2798954

Qy         1646 laAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrP 1666
                        :::      |||||||||      |||   |||      |||   |||   |
Db      2798955 ACCGCCGTCACCGCCGTCCCCGCCATCA-CTGCCATGACCGCCGTCACCGCCGTTGCCGC 2799013

Qy         1666 ro------------ProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnA 1682
                ||            |||   :::               ||||||   |||
Db      2799014 CGGCTCCGCCGAGCCCGCC-GTCACCGCCAACGCCGCCGACACCGTGGCT---------- 2799062

Qy         1682 rgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrA 1702

Db      2799062 ------------------------------------------------------------ 2799062

Qy         1702 laMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaL 1722
                              ||||||:::   ||||||      |||      :::::::::
Db      2799063 --------------GCCGACCTGACCCGCGGGTGCGGCCCCGCCGGCGCCGCCATC---- 2799104

Qy         1722 euAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeuProV 1742
                           |||||||||
Db      2799105 --------ACCGCCGGGCCC---------------------------------------- 2799116

Qy         1742 alLeuVal-ProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeu 1761
                     ||| ||||||||||||   ||||||   :::
Db      2799117 --GCCGTCACCGCCAACGCCGCCGACACCGCCGTCGCCG--------------------- 2799153

Qy         1762 ProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGly 1781
                |||      |||   |||   ::::::          |||:::|||   :::|||
Db      2799154 CCTTTCCCGCCGGCGCCTCCAACGGCCTCAGCGCTGTCGGCGCCGGAGGCGCCC------ 2799207

Qy         1782 ProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArgAsp 1801
                         |||      |||   ||||||||||||
Db      2799208 ---------TTGCCGCCGCCGCC-ACCACTAGCTCCGGC--------------------- 2799236

Qy         1802 ArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThrVal 1821
                                  |||            ::::::   |||   :::|||
Db      2799237 ------------ACCACCCGCACCGCCGGCCCCGCCCTTGCCACCGGGTCCACCGTCGCC 2799284

Qy         1822 GluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySerSer 1841
                   |||   |||                           :::|||::::::||||||:::
Db      2799285 CGACACCGACCC---------------------------ACCGGCGCCACCGGCTCCGCC 2799317

Qy         1842 GlyGlyGlyGlyGlySerSerSerArgProAlaSerHis--------------------- 1854
                      |||         :::   |||||||||   |||
Db      2799318 ------GGCACCGCCCGCGCCGC-CGGCCTGCCCCGCATCGCCTCGACCGCCGTTGCCGC 2799370

Qy         1855 ------------------SerHisAlaHisGlnHisSerProIleSerProArg------ 1866
                                  |||      |||:::|||            ||||||
Db      2799371 CACTACCTAACGCCGAACTCCCGGCACCACCGTCACCGCCGGCACCGCCCGCGCCGCCAC 2799430

Qy         1867 ThrGlnAspAlaLeuGlnGlnArg---ProSerValLeuHisAsnThrGlyMetLysGly 1885
                   |||      |||   :::|||   |||      |||               :::|||
Db      2799431 CGCCAACGCCGCCTTGACCGCCGTTGGCCTCGCTTCCTTCGCCCGGTTGCCCCGAGAGGG 2799490

Qy         1886 IleIleThrAlaValGluProSerLysProThrValLeuArgSerThrSerThrSerSer 1905
                      :::|||      ||||||   ||||||                        :::
Db      2799491 TGCCG-TCGGCG------CCGTCCTCACCTACGTTC--------------------GCG 2799522

Qy         1906 ProValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeu 1925
                |||      |||||||||      ||||||:::         |||         |||
Db      2799523 CCATTCGCACCGGCCGCGCCGCTACCACCGTCGCCGCCGGCCCCGCCGTCACCGACCAAC 2799582

Qy         1926 AspGlyValTyrProThrLeuMetGluProValLeuLeuPro---------LysGluAla 1942
                      :::   |||            |||         |||               |||
Db      2799583 CCGCCGTTACCACCGGCACCGCCCTTGCCGCCGTTCGCGCCAGCCACGGTGGCGTCGGCG 2799642

Qy         1943 ProArgValAlaArgProGlu---------------------ArgProArgAlaAspThr 1955
                |||            |||                           |||   |||
Db      2799643 CCGTTACCGCCCTTGCCGCCGTTGCCACCACTGTGCACGGTAGCGCCGGTCGCGCCGGTC 2799702

Qy         1956 GlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerPro 1975
                      :::            |||||||||               |||         |||
Db      2799703 ACACCCTCGGTGGCGCCGGCGCCGCCGGCG---------------CCACCCTTCCCACCG 2799747

Qy         1976 SerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThrIleAla 1995
                :::   ||||||   |||   |||      ||||||            :::      |||
Db      2799748 GCGCCTGGGTCACCACCATCGCCGCCGGCCCCACCG------TCACCATCCTTGAAAGCC 2799801

Qy         1996 ArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProPro 2015
                   :::|||         |||   |||         :::      |||||||||||||||
Db      2799802 ATGTCGCCGCGGCCACCGCTGCCCCCGTTACCGGGCGCC------CCACCAGCCCCGCCG 2799855

Qy         2016 AlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGlu 2035
                |||      |||   |||                        |||:::
Db      2799856 GCACCACCGTCCCCGCCA------------------------ACACCCTGG------------ 2799882

Qy         2036 LeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluPro 2055
                |||      |||                              |||   |||      |||
Db      2799883 CTGCCCGCCCGA------------------------------CCCGCAGGTGCGTCACCG 2799912

Qy         2056 ValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGlu 2075
                   :::|||   :::   ||||||                     |||
Db      2799913 CCAGCCCCACCGGCCCCACCGTCACCGCCGCGACCGCCGGCTCCGCCATCA------CCA 2799966

Qy         2076 GluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLys 2095
                   |||      |||               |||:::|||      |||   |||
Db      2799967 CCGTTGCCGCCGTCAGATACGAGCACAGCATTGAAACCGTGAGCTCCGTTACCA------ 2800020

Qy         2096 LeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSer 2115
                            |||         |||   |||   |||   |||   :::   ||||||
Db      2800021 ---------CCGGCCCCGCCGGCCCCACCGTTGCCGCCGGCACCGCCGGCCCCGCCATCG 2800071

Qy         2116 ------------------------------------SerSerProLeuLeuGlnThrAla 2123
                                                    :::   ||||||      ::::::
Db      2800072 CCGGCGTGGCCCCACCCGCGCCGCCGGCCCCACCGGCCCCGCCGTTACCTCCATCCTCA 2800131

Qy         2124 ProGlyValLysGlyHisGlnArgValValThrLeu------AlaGlnHisIleSerGlu 2141
                |||||||||            |||   :::   :::      |||
Db      2800132 CCGGGGGTACCGGATGAACCCAGGAAGATCGCCGTCATATCGGCATAGCCGGCACCCGCG 2800191

Qy         2142 ValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro--- 2160
                      :::                     |||         :::   |||||||||
Db      2800192 GCTCCGTCACCGCCATGACCGCCGGCCCCACCGTCACCGACCAACCCGCCGTTGCCACCG 2800251

Qy         2161 ---AlaProLeuTyrSerPheProGlyAlaSer---------------CysProValLeu 2174
                      ||||||         |||   |||:::                  |||   |||
Db      2800252 GCACCGCCGTTACCGCCATGACCGCCGGCGACCGGTGCGTGGGCGCCATTGCCACCTTTG 2800311

Qy         2175 AspLeuArgArgProProSerAspLeuTyrLeuPro-ProProAspHisGlyAlaProAl 2194
```

```
                           |||   :::         |||||| ||||||||||||   ||||||||
Db     2800312 CCGCCGTTGCCGCCGCTGACCGGCCCGTCGTTGCCGGCCGCCAGACCATTG-GCCCCCGC 2800370

Qy        2194 aArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThrSerVa 2214
               ||||               :::               |||   |||   :::
Db     2800371 GCGAGCACCGGCGCCGCCCGACGCCCCCCCGGCTCCGCCAGCCCCACCCAGG-------- 2800422

Qy        2214 lLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThrGluPr 2234
                                           |||   |||||||||               ||
Db     2800423 ----------------------------CCCGGGTCGCCACCGTGACCGCCGGCCCCGCC 2800454

Qy        2234 oGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluPr 2254
               |      :::   |||   :::   ||||||                           ||
Db     2800455 GTCACCGGCGGCCAGCCAACTGCCACCGTTG---------------CCGCCGGCACCGCC 2800499

Qy        2254 oSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheSerLy 2274
               ||||               ||||||               ||||||:::
Db     2800500 GTCACCAGGCGCTCCACCCAGCCCCCCACCCCCGCCAGCCCCGCCGTCTCCGGCCCGGCC 2800559

Qy        2274 sLeuThrGluSerAsnSerAla-MetValLysSerLysLysGlnGluIleAsnLysLysL 2294
                         ::::::   |||       :::      ::::::         ::::::
Db     2800560 GACATACCCCAATAGTCCAGCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGTT 2800619

Qy        2294 euAsnThrHis------------------------AsnArgAsnGluProGluTyrAsnI 2306
                       |||                        :::|||      |||   |||
Db     2800620 TCCGCCGGCACCGCCATTGCCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGG- 2800678

Qy        2306 leSerGlnProGlyThrGluIlePhe-AsnMetProAlaIleThrGlyThrGlyLeuMet 2325
                                   :::|||       |||         ||||||
Db     2800679 -----------CCGGCCCATTTATTCCGGTAGCCCCGGCACCGCCGGCACCGCCGGCCGC 2800727

Qy        2326 ThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIle 2345
               |||      :::                     :::|||         :::
Db     2800728 ACCGGCACCACCGGCCCCGCCGACACCGCCAACGCCACCGGCGCCGCCGTTACC------ 2800781

Qy        2346 IleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAla 2365
                                                   |||:::   |||      :::
Db     2800782 ------------------------------------GAGAAGCCACCCTCCCCGACCACC 2800805

Qy        2366 AsnAlaPheAsnProLeu---AsnAlaSerAlaSerLeuProAlaAlaMetProIleThr 2384
                  |||   :::|||         |||   |||   :::               :::|||
Db     2800806 GTTGCCGCCCACCCCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACCGTCACC 2800865

Qy        2385 AlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLys 2404
               |||                     |||:::      ||||||         :::
Db     2800866 GCCGGC------------------ACCGTCGTCCGACCCGGCAACACCAGCCGCCCCATC 2800907

Qy        2405 ValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGly 2424
               ::::::||||||   :::            ::::::               ||||||
Db     2800908 CTGACCGGGCGTAGCACCGTTGGCCCCGGCCGCACCTACACCACCCACGCCGCCAGCGCC 2800967

Qy        2425 ------AspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArgArg 2442
                        :::||||||   |||::::::      :::   |||
Db     2800968 CCCATGACCAAACCACCCCGCGTTACCGCC---CGCGCCGCCGGCGCCACCACCAGCCCC 2801024

Qy        2443 ThrProLeuThrAsnArgVal-TrpGluAspArgProSerSerAlaGlySerThrProPh 2462
               |||      |||                :::||||||:::         :::   |||
Db     2801025 ACCGGCACCACCGGCGCCGCCGTTGCCGAACAGGCCCGCGTTGCCACCCGCCCCGCCGGC 2801084

Qy        2462 ePro------------------------------------------TyrAsnProLeuIl 2468
                |||                                             :::|||   ::
Db     2801085 GCCACCACCAGCCCCACCCATACCGCCGACGCCGCCCCCACCCGCCAGCCATCCACCCGT 2801144

Qy        2468 eMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySe 2488
               :      :::      |||      |||   |||   ||||||      |||      ::
Db     2801145 CCCGCCGGTACCTCCGGGTGCGCCCGCCCCGCCCGCGCCACCGGCGCCGCCGATCCCAAA 2801204

Qy        2488 rGlyProLeuAlaGlyPro 2494
               :   |||   |||   |||
Db     2801205 TAACCCGGCCGCCCCGCCG 2801223

RESULT 13
US-09-103-840A-1
```

```
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:               1.28e-18       Length:         4411529
Score:                   665.00         Matches:        627
Percent Similarity:      33.04%         Conservative:   296
Best Local Similarity:   22.44%         Mismatches:     1111
Query Match:             5.03%          Indels:         776
DB:                      3              Gaps:           122

US-09-522-753-5 (1-2517) x US-09-103-840A-1 (1-4411529)

Qy          20 ProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeu 39
               |||||||||::::::|||   |||   |||      |||                  :::
Db     2798750 CCCCCGCACGCGGTCAGCGC-CCGCGCCAGCAGGGCCGG------------------ATC 2798790

Qy          40 LeuGluTyrGlnHisHisSerArg------------------------------AspTyr 49
               |||:::         :::|||:::                              |||
Db     2798791 CTCGATGTCGCGCGCTATAGCCAGCGCCTTCTGGGCATCATCTAGGCGGTCGGTGATGCC 2798850

Qy          50 AlaSerHisLeuSerProGlySerIle---------------IleGlnProGlnArgArg 64
                        :::      |||   :::               |||||||||:::   |||
Db     2798851 GGCCACGGCATCTATCAGGGCCCGGTCGGCCAGTGCCCGCGCATACAACCCAGGGT-CGG 2798909

Qy          65 ArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGlnGluLeuHisLeu 84
                  |||      :::      .  :::||||||      |||::::::         |||
Db     2798910 CCCCCGCCGGATGTGCATCGTGGTCGGCCAGGGCGGCGGCGAACCAGGCCAGCCCCTCTT 2798969

Qy          85 ------------------ArgProGluSerHisSerTyrLeuProGluLeuGlyLysSer 98
                                          :::         :::|||         :::
Db     2798970 GCAGGCGGCCCCGGGCACGCCACAACGGCTGCAGACATGATGCCAACAGCAACGCGTGGC 2799029

Qy          99 GluMetGluPheIleGluSer---------LysArgPro---------ArgLeuGluLeu 112
                                             |||||||||         |||
Db     2799030 CGGTATCGCCATTCTCGCGGCTGAACGCGAAAGCGGCCCGTAGGTTGTCGATCTCGA--- 2799086

Qy         113 LeuProAspProLeuLeuArgProSerProLeuLeuAlaThrGlyGlnProAlaGlySer 132
                                 ||||||                  |||:::|||
Db     2799087 ---------------GCTCGGCCTGGTTGAGCCGGCGTTCATGGCCGGCCACCGAGGGGG 2799131

Qy         133 GluAspLeuThrLysAspArgSerLeuThrGly------------------------Lys 144
                           :::   |||      :::|||                        :::
Db     2799132 CGTCAAGCCCGGCGGCAACGGCCGCGTAGTGGTCGCGGTGTCGCGCACGCACGGCATCGG 2799191

Qy         145 LeuGluProValSerProProSerProProHisThrAsp---ProGluLeuGluLeuVal 163
                        :::         ||||||   ||||||      |||   |||
Db     2799192 CATCGCCGGATTCACGCAGCTTCTCCAACGCATACTGGCGCACCGTCTCTAGCAGGCGGT 2799251

Qy         164 ProPro---------------ArgLeuSerLysGluGluLeuIleGlnAsnMetAspArg 178
                                    |||   |||      |||
Db     2799252 AGCGCGTTCGGCCGTCGCTGTCGTCGGTCACCACCAGAGACTTGTCTGCCAGCAGGCTGA 2799311

Qy         179 ValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGln 198
                  ||||||               |||:::         :::|||            :::
```

```
Db     2799312 GCAGATCGACCACCTCGTAGCGCTGAACGTCACCGCCGGCGGCTGCCGCTTGGGCACCGT 2799371

Qy         199 GlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSerProProPro 218
               :::   :::      :::   |||||||||      |||         :::|||   |||
Db     2799372 CGAGATCAAACCCGCTCGGGAAAACCGCCAGTCGCCGAA------ACAGCACCTGCTCCG 2799425

Qy         219 IleGluSerLysHisArgSerLeuValGlnIleIle-TyrAspGluAsnArgLysLysAl 238
               :::            ||||||         :::       |||   :::      :::
Db     2799426 GTC------------CGGTCAGCAGCGCATGTGACCAGTCGACGGAAGCCCGCATCGTCT 2799473

Qy         238 aGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAs 258
                   ||||||||||||               |||   |||   |||   |||:::
Db     2799474 GCTGGCGGCGCACCGCAATACGCGATCCACCGGTCAGCAGGCGGAA---CCGGTCATGCA 2799530

Qy         258 nGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLy 278
                         |||      |||
Db     2799531 AGCTGTCGACGATTTCGGTCAG-------------------------------------- 2799552

Qy         278 sLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPh 298
                                  ::::::      |||   |||:::
Db     2799553 -------------------CGCCAGGGCACGCACCCGCGACGC----------------- 2799576

Qy         298 eCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGlu----------- 314
                |||   ::::::|||:::         |||:::|||         :::
Db     2799577 -TGC---AAGTTCGATCGCCAGCGGAATGCCGTCGAGTCGGTGGCAGATCTCGGTCACCA 2799632

Qy         315 -------------------ArgIleGluAsnAsnProArg-------------------- 321
                                     :::|||      ||||||
Db     2799633 GGGCGAGGTTGTCGGCAGTGATCTCGAGTTCGGGCCGCGCCTCACGAGCGCGGTCGGTGA 2799692

Qy         322 -------------------------ArgArgAlaLysGluSerLysValArgGluTyrTy 333
                                        :::|||               :::         ::
Db     2799693 ACAACTCGATCGCCTCGCCGTGCCCAGCGGGGAACCCGCCAAATCTGCTCACCGGCCA 2799752

Qy         333 rGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerAr 353
               :      :::||||||      |||:::      :::         |||   :::   ||
Db     2799753 CCGCGATCGGTTCCCGGCTGGTCGCCAATACCCTCAGCGCTGGGCACGCCCCGAGCAACG 2799812

Qy         353 gValGlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHis---GluVa 372
               |      |||      :::   :::            |||      :::|||   :::
Db     2799813 CGACGATCAGAGCCGCGCACCCGTCGAGCAAGTGCTCGCAGTTGTCCAGCACTACCAGCA 2799872

Qy         372 lSerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLe 392
                :::   :::   |||      :::            :::   ::::::      :::||
Db     2799873 TGCGCCGGTCGCCGAT------ACGCCGCACAATGGTGTCCACCGTCGAGCGGCCCGGCT 2799926

Qy         392 uAlaValIlePro------ProMetLeuTyrAspAlaAspGlnGlnArgIle-------- 407
               |         |||      |||         |||      |||:::|||:::
Db     2799927 GATCCGGCAACCCCAAAACCCGCGCCGCCGCGATCGGCACCAGCGCCGGGTCGGTGATCG 2799986

Qy         408 ----------------------------------------------------LysPheIl 410
                                                                   :::   ::
Db     2799987 GCGCCAGGTTGACATACCAAACCCCGTCCGGATAACCGTCGGCAACGGCGCTCGCGACCT 2800046

Qy         410 eAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAs 430
               :               ::::::||||||            |||               ::
Db     2800047 GTGTCGCCAGGCGTGTCTTTCCGACCCCGCCGACACCGGTAAGGGTGACCCACCGTTTGA 2800106

Qy         430 nMetTrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAs 450
               ::::            ::::::   |||            |||:::      |||
Db     2800107 CGTC------------CAGCAGCCCACGGACTTGCGCCACTTCGTCGACGCGCCCCACCA 2800154

Qy         450 nPheGlyLeuIleAlaSerPheLeuGluArgLysThrVal-------------------- 463
                               |||   |||   ||||||||||||
Db     2800155 G------------CCGAGTGAGCTGGGCCGGAAGACAGTGCGCACCAACGACTTTCCGGG 2800202

Qy         464 ----------AlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLy 480
                            |||   |||:::         :::|||:::   :::      :::::
Db     2800203 TCCGCAGCGGCGGGAACGCGTTGTGCAGATCAGGGTGACACAGCTGCACCACCCGTTCCG 2800262

Qy         480 sSer------------------LeuValArgArgSerTyrArgArgArgGlyLysSerGl 494
               :|||                  :::               :::   |||         ||
Db     2800263 GTCGGGGCAGGTCGTCCAGCCGGTAGGTACCGAGGTCGTTCAGCCACGCGTCCTTGGGCA 2800322
```

```
Qy         494 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProAr 514
               ||||         |||               :::|||            |||:::   ::
Db     2800323 GCAGGTCAGCAACCAGATCGCTGGTAGTTCCCGACAACACGGTCTGGCCCCCGTGGGCCA 2800382

Qy         514 gSerSerGln-------------------------------------------------- 517
               :   :::|||
Db     2800383 GCTCGCGCAGCCGGGCGGTGCGGTCGATGGTCGGCCCTACGCAGTTGCCCTCGTCGGGTG 2800442

Qy         518 -GluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysPro-- 536
                                        :::   :::   ::::::   :::   :::|||
Db     2800443 ACGACACCTCCCCGGTGTGCATGCCGATGCGCAGCCGGATCGGTGCCAGCGGCGCCCGCT 2800502

Qy         537 -GluValGluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGl 556
                :::   :::      :::      :::                        :::|||||
Db     2800503 GCAAGCCCAGGGCGCACGCCACGGCGTC------------------GGATGCGCGGGCGA 2800544

Qy         556 uAspAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThr------------Al 572
               |   :::   |||      |||   |||         |||                  ||
Db     2800545 ACGCCACCAAGAAGCTGTCGCCTTCGCCCTGTTCGACCGGGCAAACCCCGCGGTGCTCGC 2800604

Qy         572 aAsnSerGlnGlyArgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSe 592
               ||||      |||:::|||   |||   :::   ||||||   :::   :::|||   ::
Db     2800605 GAACCAATTCGGTCAGCGTTC-GGTCCAGTTTGGCGATCGCCGTCGTGTCAAGCTGAGAC 2800663

Qy         592 r------------------------------GluGluAlaIleThrProGlnGlnSerAl 602
               :                                    |||   :::|||      |||
Db     2800664 CCCGGCAGGTGGGTCGCGCCCTCGATATCGGCCAGCAGCAACGTCACCGTGCCCGTCGGT 2800723

Qy         602 aGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGlu----------- 618
                :::      |||      |||      |||   :::   |||
Db     2800724 ACAAGCTCGCTCACACCATCTGCGCTCCAGTCCACAGGT-ACCACGTCGACGCCGGGGTG 2800782

Qy         619 -------------------GluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyAr 632
                                     :::|||   |||:::      |||      |||
Db     2800783 AATCTTGCTCATGCTAGCCAGCATCGAGCCAGCGCGTAGCGCATTACATCGGCAC----- 2800837

Qy         632 gAsnTrpSerAlaIleAlaArgMetValGlySerLysThrVal---------SerGlnCy 649
                                  |||:::      ||||||::::::         |||   ||
Db     2800838 ----------------CTGCGCCTAGATTGCTCGAAATCTCTTGGCCGCCGGTCCATGTG 2800881

Qy         649 sLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGlu-------------- 664
               |         :::            :::|||
Db     2800882 TTCTACGCGCTTTAGTCGATGCATTCGGCGACCGGCGTGCCATCGCGGCGGACCTACAGT 2800941

Qy         665 ----IleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArgArgLysLy 683
                   :::|||               :::               :::      ||||||
Db     2800942 GCCCGTGCTGTCCGCTGGCAATTGTGAGTCCCCCAGTGCTGGCAGCATCGCCCGCAAGAA 2801001

Qy         683 sLys----------------------------LysAlaProAlaAlaAlaSerGluGluA 694
                :::                            |||:::||||||      |||   |||:
Db     2801002 CCGACACGACCGCATCGTGGGCGGTGCCGTCGAAGTCGCCGGCTGACCGATCGGCGGAGT 2801061

Qy         694 la------------AlaPheProProValValGluAspGluGluMetGluAlaSerGlyV 710
               ::               |||      :::      |||::::::      |||:::
Db     2801062 CACCGGCCCGATGGGGTTTCCGAAGGCTAGGGAATGATGACGATGGGGCGGCCGCCTCGG 2801121

Qy         710 al------------------------SerGlyAsnGluGluGluMetValGluGluAlaG 722
                                         :::   ||||||      :::
Db     2801122 CCGCCTTCGCCGTAACCCCCAACCATGCGGAAAACGAGCCTAGCGTCGCCCGGCCGCGCA 2801181

Qy         722 luAlaLeuHisAlaSerGly---AsnGluValProArg------------------GlyG 735
                 |||   |||      |||   |||:::      |||                  |||
Db     2801182 GAGCGAGCCATCGCGGTGGCGCCAACGACAGGAAGCGATCCGGATTCTCTGACCATGGTG 2801241

Qy         735 luCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProH 755
                 |||||||||      |||:::         ::::::   |||      |||||||||
Db     2801242 GGTGTTCTGGCTACGTGACGTTA---------ACGGAGATGGAGGGGCCGCCTTCGCCG- 2801291

Qy         755 isThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyA 775
                                                  |||   ||||||
Db     2801292 -----------------------------------CCTTCACCGCCG------------- 2801303
```

```
Qy        775 laAspGlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleG 795
                :::   |||   ||||||   |||   |||                  :::|||   |
Db    2801304 --GAACCGCCGGAGCCAGGGTCGCCCCTCCCGTTGCCGGAGCCACCCGACTCGCCCGACG 2801361

Qy        795 luProThrProAlaSerGluAlaThrGlyAlaProThrPro------ProProAlaProP 813
              |||||||||||                  ||||||   |||      ||||||   ||||
Db    2801362 AGCCGACGCCGCCGGAGGTCAAGCCACCGGCACCGCGTCCGCCGTCACCTCCGCGCCCGC 2801421

Qy        813 roSerProSerAlaProProProValValProLysGluGluLysGluGluGluThrAlaA 833
              ||||||||   :::|||||||||::::::   :::            |||         |
Db    2801422 CGTCCCCGCCGTCACCGCCGCCGATGCTGCG-AGGCGGAGGGGCGCCGAAGCCGCCGGAG 2801480

Qy        833 laAlaProPro------------ValGluGluGlyGluGluGlnLysProProAlaAlaG 849
              |||||   |||            |||      |||               |||
Db    2801481 CCGCCGGTCCCGCCGTCGCCTCCGTCACCACCGGGGGCGCCACCGTCTCGCCCGGCCCCA 2801540

Qy        849 luGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluG 869
                :::   ||||||            |||      |||      :::      ::::::
Db    2801541 CCCAAGCCGCCGTT---GCCGCCGTTGCCACCCGGCCCGCCGTCGCCTGCATCAGCAAAG 2801597

Qy        869 luAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluG 889
                |||         |||                        |||      ||||||   |
Db    2801598 CTGCCGTTGCCGTCCCCACCGTGACCGCCGTTCCCGCCGTCGCCTCCGTCACCGCCGGGG 2801657

Qy        889 lyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThr-----------T 905
              |||||               :::   |||   :::   |||||||||            |
Db    2801658 GCGCC---------------GAAGCCGGCCTTGCCGTGCGCGCCACTTGTGGAACCGAAA 2801702

Qy        905 hrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspG 925
              |||||   |||:::||||||         :::   ::::::   |||      |||
Db    2801703 CCGCCTTGTCCGCCGGGGCGGCCCCACCCGCCGTCGCCGCCGTCACCTCCGTCGCCGCCA 2801762

Qy        925 luValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer-ProArgProSerLeu 944
                :::         :::      :::      :::   :::::: |||:::|||   |||
Db    2801763 GGCTC---TCCGTCAAAATCCGCGAGATAGGTAAAGCCGTCACCGCCCAAGCCACCATTA 2801819

Qy        945 LeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLys 964
                             ||||||   :::      ||||||      |||         :::
Db    2801820 CCAGCGTCCCCGCCCGACCCGCCGTCACCGCCGTCCCCGCCAACGCCT---------CGA 2801870

Qy        965 GlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluPro 984
                             |||   |||   ||||||                        |||
Db    2801871 TTGCCGACCTCGCCGGCGGGTGCCGACCCGCCG---------------------GCCCCG 2801909

Qy        985 ProArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsn 1004
              |||         |||   |||      |||   ||||||:::||||||   |||   :::
Db    2801910 CCGTTTCCGCCGGCGCCGCCCCACCCGCCGTAGCCACCGTCGCCGCCGTCGCCGCCGTCG 2801969

Qy       1005 ---------------------------------------LeuGlnProGluSerAspAla 1011
                                                     |||   |||         |||
Db    2801970 CGGCCCGTCGTTTCGTTAATGTCAAAGCCGTCAACGCCGTTACCGCCGACCCCACCAGCC 2802029

Qy       1012 ProGlnGlnProGlySerSerProArgGlyLysSerArg---------SerProAlaPro 1028
              |||   :::|||   :::   |||         |||               |||||||||
Db    2802030 CCGCCTAGGCCTCCGGCCCCGCCGTCACCACCGTCGCCGGTCTGAGTTCCGCCGGCGCCA 2802089

Qy       1029 ProAla---------------------------------------AspLysGluAlaPhe 1035
              ||||||                                             :::   |||
Db    2802090 CCGGCCCCGCCGTCGCCTCCCGCCCCACCGCTGCCGCCGTCGAAGCCGTCGAAGCCCTTT 2802149

Qy       1036 AlaAlaGluAlaGlnLysLeuProGlyAspPro------------ProCysTrpThrSer 1051
                 :::|||:::         |||      |||            |||   |||
Db    2802150 AGGTCGGAGTCGGGCGACCAACCCGCGCCACCGGCCGCGCCGTTGCCTCCCTGGCCG--- 2802206

Qy       1052 GlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAsp 1071
                          ||||||||||||         |||   |||||||||   ||||||
Db    2802207 ------------CCGGTTCCGCCGCCGCCGTTCATCCCGGCGTCGCCGCCCGCCCCGCCG 2802254

Qy       1072 ProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThr 1091
                                   ||||||      |||   |||   ||||||
Db    2802255 ------------------TGTCCACCAACCCCGCCGCCGCCGCCGGGGCTG--------- 2802287

Qy       1092 AlaArgProValLeuProArgProProThrIleSerAsnProProPro-LeuIleSerSe 1111
                    |||      |||   ||||||               |||||| |||      ||
Db    2802288 ------CCGCCCCGGCCAGCCCCACCTTGG-----------CCGCCGGCTCCGCCGTTC 2802329

Qy       1111 rAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerVa 1131
              |   :::   |||   :::|||   |||:::
Db    2802330 CCGCCGTCGCCCAGAAATGCTCCGCCGGCGC----------------------------- 2802360

Qy       1131 lGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGl 1151
                     |||   |||:::      |||         |||:::   |||   |||:::
Db    2802361 -------CACCAGCCCCACCGGCGCCACCA-GCCCCACCGTTGCCGCCAGCAACGGTGAG 2802412

Qy       1151 yLeuPro---LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
                  |||      |||      |||            |||   |||
Db    2802413 CCCTCCGAGGGCACCGTGCGCGCCGTCGCCACCCTTGCCGCCGTCA-------------- 2802458

Qy       1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
                     ||||||         |||   ||||||            :::|||   |||
Db    2802459 -CCGCCGTCACCGATGTCGCCGGCGTCACCGCCC------------TTGCCTCCAGCCCC 2802505

Qy       1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
                  |||      :::         :::   ||||||         |||   ||||||
Db    2802506 ACCGGCCCC-GCCATCACCGCCGAGAGCTTCGGCAGCGGTGCCGTCGGCCCCATCACCAC 2802564

Qy       1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
               |||                  |||      :::|||:::|||         ::::::
Db    2802565 CGGCTCCGCCGTCCCCGAATAGCCCGGCGTTGCCGCCGTCACCGCCCTGGCCGCCGTCGC 2802624

Qy       1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
                  |||                        ||||||:::
Db    2802625 CGCCGGCCGCGGCGGC------------CTTGGCACCGTTGCC----------------- 2802655

Qy       1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
               :::   :::   :::      |||:::                     |||   :::
Db    2802656 -GCCGACGCCGCCGTCGCCGCCGGTCAGTGGCCCGTGTTTGCTGGCGTCCACGCCGTT-- 2802712

Qy       1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
                  |||:::   |||   |||   :::               :::   |||   |||::
Db    2802713 ----GGCCGCGGAGGTGCCGTTGCCGCTGTCACCCCCCAGACCGCCGCGACCGCCTGCGC 2802768

Qy       1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
              :                                       ||||||         |||::
Db    2802769 C-------------------------------GGGGTCACCGCCGTTACCGCCCGCTC 2802795

Qy       1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
              :            :::|||            |||:::   :::|||
Db    2802796 CGCCGGCGCCGCCGACGGTGATACCAATGCCGCCGTT---GCCGCCGGCCCCGCCAACGC 2802852

Qy       1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
                  |||   |||            ||||||   ||||||      ::::::
Db    2802853 CGCCGGCGCCGCCGA-GTCCGCCGTCGCCACCGACCCCACCGGTGCCGTGACTGCCGA-- 2802909

Qy       1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
                                          |||         |||      :::      ::
Db    2802910 ---------------------------CCGTCCCCGAAGGTGCGGCCCCGCCGACCCCA 2802941

Qy       1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
              :      |||               :::|||            |||||||||||||||::
Db    2802942 CCGTCCCCGCCATGTCCACCGACCCCGCCGGCACCG-CCATCGCCGCCACCACCGCCGGC 2803000

Qy       1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
              :      :::                        |||   ||||||
Db    2803001 CCCACCGGTGCCGCCGATACTGTCGATA---CCGTTGGCGCCCCTG-------------- 2803043

Qy       1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
                                                                     |||
Db    2803044 ----------------------------------------------------GCCCCGGC 2803051

Qy       1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGly-- 1449
                     |||      ||||||   ||||||   ||||||   ||||||      |||
Db    2803052 CCCACCGCTAGCGCCCACACCGCCGTTGCCGCCGGCCCCG---CCGTTGCC-GCCGGCAC 2803107

Qy       1450 -SerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLy 1469
               ::::::|||:::      |||            |||||||||:::
```

```
Db     2803108 CGCCGTCACCCGACACCGACCC------------ACCGGCGCCACC-------------- 2803141

Qy        1469 sLysHisAspValArgSerLeuIleGlySerProGlyArgThr-PheProProValHisP 1489
                                        |||:::   |||   |||    ||||||      |
Db     2803142 -------------------------GGCACCACCGGCACCACCGGCACCGCCGGCCTGCC 2803176

Qy        1489 roLeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuL 1509
               ||
Db     2803177 CC---------------------------------------------------------- 2803178

Qy        1509 ysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValI 1529
                 :::   |||                  |||         ||||||   :::
Db     2803179 --GCGTCGCCCTGACCCCCGTTGCCTCCCGGTTGGCCGAGGGCGAGGGCATCT------- 2803229

Qy        1529 leValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaP 1549
                       |||   |||   |||      :::||||||                     |
Db     2803230 --------GAACCAGGCGCGCCCGAATTGGCCCCGTTGGCGCCGGCCGCGCCATCGCCAC 2803281

Qy        1549 roPheAlaGlyHisLeuProArgGlySerProValThrMetArg---------GluProT 1566
               ||            ||||||         |||         |||            |||:
Db     2803282 CA------------TTGCCGCCGGCGCCACCATCGCCGACCCGGCCGGCATTGCCGCCGT 2803329

Qy        1566 hrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuT 1586
               ::|||   |||
Db     2803330 CGCCTCCGTTGCCC---------------------------------------------C 2803344

Qy        1586 hrSerThrProArgGluIleAlaLysSerProHisSerThrValProGluHisHisProH 1606
                 :::   |||      :::|||   :::|||         :::|||         |||
Db     2803345 CGGCGCCGCCCGCGACGCTGGCTTGCGCACCGTTGCCACCGTTACCACCGTTGCCGCCGC 2803404

Qy        1606 isProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgS 1626
                 |||      |||:::                  ::::::         |||      :
Db     2803405 TGCCG---GGCCCGTGG---------------TCGCTGGCGTCCACACCGCTG---GCCG 2803443

Qy        1626 erHisIleProLeuAlaPheAspProThrSerIlePro-ArgGlyIleProLeuAspAla 1645
               ::   :::||||||         |||         ||| ||||||:::|||
Db     2803444 CGCGGGTGCCGTTGCCGCTGTCGCCGCCCAAG---CCGCCGAGGCCTCCCGCGCCGGGGT 2803500

Qy        1646 AlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyr 1665
                        :::      |||||||||      |||   |||      |||   |||
Db     2803501 CACCGCCGTCACCGCCGTCCCCGCCATCA-CTGCCATGACCGCCGTCACCGCCGTTGCCG 2803559

Qy        1666 Pro------------ProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsn 1681
               |||            |||   :::               ||||||   |||
Db     2803560 CCGGCTCCGCCGAGCCCGCC-GTCACCGCCAACGCCGCCGACACCGTGGCT--------- 2803609

Qy        1682 ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThr 1701

Db     2803609 ------------------------------------------------------------ 2803609

Qy        1702 AlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeuAla 1721
                              ||||||:::   ||||||      |||      :::::::::
Db     2803610 ---------------GCCGACCTGACCCGCGGGTGCGGCCCCGCCGGCGCCGCCATC--- 2803651

Qy        1722 LeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeuPro 1741
                           |||||||||
Db     2803652 ---------ACCGCCGGGCCC--------------------------------------- 2803663

Qy        1742 ValLeuVal-ProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLe 1761
                     ||| ||||||||||||   ||||||   :::
Db     2803664 ---GCCGTCACCGCCAACGCCGCCGACACCGCCGTCGCCG-------------------- 2803700

Qy        1761 uProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGl 1781
                |||      |||   |||   ::::::         |||:::|||   :::|||
Db     2803701 -CCTTTCCCGCCGGCGCCTCCAACGGCCTCAGCGCTGTCGGCGCCGGAGGCGCCC----- 2803754

Qy        1781 yProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArgAs 1801
                         |||      |||   ||||||||||||
Db     2803755 ----------TTGCCGCCGCCGCC-ACCACTAGCTCCGGC-------------------- 2803783

Qy        1801 pArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThrVa 1821
                                  |||            ::::::   |||   :::|||
Db     2803784 -------------ACCACCCGCACCGCCGGCCCCGCCCTTGCCACCGGGTCCACCGTCGC 2803830

Qy        1821 lGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySerSe 1841
                   |||   |||                           :::|||::::::||||||::
Db     2803831 CCGACACCGACCC---------------------------ACCGGCGCCACCGGCTCCGC 2803863

Qy        1841 rGlyGlyGlyGlyGlySerSerSerArgProAlaSerHis-------------------- 1854
               :      |||         :::   |||||||||   |||
Db     2803864 C------GGCACCGCCCGCGCCGC-CGGCCTGCCCCGCATCGCCTCGACCGCCGTTGCCG 2803916

Qy        1855 -------------------SerHisAlaHisGlnHisSerProIleSerProArg----- 1866
                                  |||      |||:::|||            ||||||
Db     2803917 CCACTACCTAACGCCGAACTCCCGGCACCACCGTCACCGCCGGCACCGCCCGCGCCGCCA 2803976

Qy        1867 -ThrGlnAspAlaLeuGlnGlnArg---ProSerValLeuHisAsnThrGlyMetLysGl 1885
                   |||      |||   :::|||   |||      |||               :::||
Db     2803977 CCGCCAACGCCGCCTTGACCGCCGTTGGCCTCGCTTCCTTCGCCCGGTTGCCCCGAGAGG 2804036

Qy        1885 yIleIleThrAlaValGluProSerLysProThrValLeuArgSerThrSerThrSerSe 1905
               |      :::|||      ||||||   ||||||                        ::
Db     2804037 GTGCCG-TCGGCG------CCGTCCTCACCTACGTTC---------------------GC 2804068

Qy        1905 rProValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLe 1925
               :|||      |||||||||      ||||||:::         |||         |||
Db     2804069 GCCATTCGCACCGGCCGCGCCGCTACCACCGTCGCCGCCGGCCCCGCCGTCACCGACCAA 2804128

Qy        1925 uAspGlyValTyrProThrLeuMetGluProValLeuLeuPro---------LysGluAl 1942
                      :::   |||            |||         |||               ||
Db     2804129 CCCGCCGTTACCACCGGCACCGCCCTTGCCGCCGTTCGCGCCAGCCACGGTGGCGTCGGC 2804188

Qy        1942 aProArgValAlaArgProGlu---------------------ArgProArgAlaAspTh 1955
               ||||            |||                           |||   |||
Db     2804189 GCCGTTACCGCCCTTGCCGCCGTTGCCACCACTGTGCACGGTAGCGCCGGTCGCGCCGGT 2804248

Qy        1955 rGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerPr 1975
                      :::            |||||||||               |||         ||
Db     2804249 CACACCCTCGGTGGCGCCGGCGCCGCCGGCG---------------CCACCCTTCCCACC 2804293

Qy        1975 oSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThrIleAl 1995
               |:::   ||||||   |||   |||      ||||||            :::      ||
Db     2804294 GGCGCCTGGGTCACCACCATCGCCGCCGGCCCCACCG------TCACCATCCTTGAAAGC 2804347

Qy        1995 aArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProPr 2015
               |   :::|||         |||   |||         :::      ||||||||||||||
Db     2804348 CATGTCGCCGCGGCCACCGCTGCCCCCGTTACCGGGCGCC------CCACCAGCCCCGCC 2804401

Qy        2015 oAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGl 2035
               ||||      |||   |||                        |||:::
Db     2804402 GGCACCACCGTCCCCGCCA---------------------ACACCCTGG----------- 2804429

Qy        2035 uLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluPr 2055
                |||      |||                              |||   |||      ||
Db     2804430 -CTGCCCGCCCGA------------------------------CCCGCAGGTGCGTCACC 2804458

Qy        2055 oValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGl 2075
               |   :::|||   :::   ||||||                     |||
Db     2804459 GCCAGCCCCACCGGCCCCACCGTCACCGCCGCGACCGCCGGCTCCGCCATCA------CC 2804512

Qy        2075 uGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLy 2095
                   |||      |||               |||:::|||      |||   |||
Db     2804513 ACCGTTGCCGCCGTCAGATACGAGCACAGCATTGAAACCGTGAGCTCCGTTACCA----- 2804567

Qy        2095 sLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSe 2115
                            |||         |||   |||   |||   |||   :::   |||||
Db     2804568 ----------CCGGCCCCGCCGGCCCCACCGTTGCCGCCGGCACCGCCGGCCCCGCCATC 2804617

Qy        2115 r------------------------------------SerSerProLeuLeuGlnThrAl 2123
               |                                    :::   ||||||      :::::
Db     2804618 GCCGGCGTGGGCCCCACCCGCGCCGCCGGCCCCACCGGCCCCGCCGTTACCTCCATCCTC 2804677

Qy        2123 aProGlyValLysGlyHisGlnArgValValThrLeu------AlaGlnHisIleSerGl 2141
               :|||||||||            |||   :::   :::      |||
Db     2804678 ACCGGGGGTACCGGATGAACCCAGGAAGATCGCCGTCATATCGGCATAGCCGGCACCCGC 2804737
```

```
Qy       2141 uValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro-- 2160
                     :::                     |||         :::   |||||||||
Db    2804738 GGCTCCGTCACCGCCATGACCGCCGGCCCCACCGTCACCGACCAACCCGCCGTTGCCACC 2804797

Qy       2161 ----AlaProLeuTyrSerPheProGlyAlaSer---------------CysProValLe 2174
                     ||||||         |||   |||:::                  |||   ||
Db    2804798 GGCACCGCCGTTACCGCCATGACCGCCGGCGACCGGTGCGTGGGCGCCATTGCCACCTTT 2804857

Qy       2174 uAspLeuArgArgProProSerAspLeuTyrLeuPro-ProProAspHisGlyAlaProA 2194
              |            |||   :::         |||||| |||||||||||    |||||||
Db    2804858 GCCGCCGTTGCCGCCGCTGACCGGCCCGTCGTTGCCGGCCGCCAGACCATTG-GCCCCCG 2804916

Qy       2194 laArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThrSerV 2214
              |||||               :::               |||   |||   :::
Db    2804917 CGCGAGCACCGGCGCCGCCCGACGCCCCCCCGGCTCCGCCAGCCCCACCCAGG------- 2804969

Qy       2214 alLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThrGluP 2234
                                           |||   |||||||||               |
Db    2804970 -----------------------------CCCGGGTCGCCACCGTGACCGCCGGCCCCGC 2805000

Qy       2234 roGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluP 2254
              ||      :::   |||   :::   ||||||                           |
Db    2805001 CGTCACCGGCGGCCAGCCAACTGCCACCGTTG---------------CCGCCGGCACCGC 2805045

Qy       2254 roSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheSerL 2274
              |||||               ||||||               ||||||:::
Db    2805046 CGTCACCAGGCGCTCCACCCAGCCCCCCACCCCCGCCAGCCCCGCCGTCTCCGGCCCGGC 2805105

Qy       2274 ysLeuThrGluSerAsnSerAla-MetValLysSerLysLysGlnGluIleAsnLysLys 2293
                         ::::::   |||       :::      ::::::         ::::::
Db    2805106 CGACATACCCCAATAGTCCAGCCGACCCGCCAGCACCCCCGGCGCCGCCCACGCCGCCGT 2805165

Qy       2294 LeuAsnThrHis------------------------AsnArgAsnGluProGluTyrAsn 2305
                       |||                        :::|||      |||   |||
Db    2805166 TTCCGCCGGCACCGCCATTGCCGCCGGCGCCGCCGTCACCGCCGGCCGCCGAGATACCGG 2805225

Qy       2306 IleSerGlnProGlyThrGluIlePhe-AsnMetProAlaIleThrGlyThrGlyLeuMe 2325
                                   :::|||       |||         ||||||
Db    2805226 ------------CCGGCCCATTTATTCCGGTAGCCCCGGCACCGCCGGCACCGCCGGCCG 2805273

Qy       2325 tThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIl 2345
               |||      :::                     :::|||         :::
Db    2805274 CACCGGCACCACCGGCCCCGCCGACACCGCCAACGCCACCGGCGCCGCCGTTACC----- 2805328

Qy       2345 eIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAl 2365
                                                   |||:::   |||      :::
Db    2805329 -------------------------------------GAGAAGCCACCCTCCCCGACCAC 2805351

Qy       2365 aAsnAlaPheAsnProLeu---AsnAlaSerAlaSerLeuProAlaAlaMetProIleTh 2384
                  |||   :::|||         |||   |||   :::               :::||
Db    2805352 CGTTGCCGCCCACCCCACCGGCACCGCCATCGCCCCCGTCCGACCCTGCCAAACCGTCAC 2805411

Qy       2384 rAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLy 2404
              ||||                     |||:::      ||||||         :::
Db    2805412 CGCCGGC------------------ACCGTCGTCCGACCCGGCAACACCAGCCGCCCCAT 2805453

Qy       2404 sValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGl 2424
               ::::::||||||   :::            ::::::               ||||||
Db    2805454 CCTGACCGGGCGTAGCACCGTTGGCCCCGGCCGCACCTACACCACCCACGCCGCCAGCGC 2805513

Qy       2424 y------AspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArgAr 2442
                        :::||||||   |||::::::      :::   |||
Db    2805514 CCCCATGACCAAACCACCCCGCGTTACCGCC---CGCGCCGCCGGCGCCACCACCAGCCC 2805570

Qy       2442 gThrProLeuThrAsnArgVal-TrpGluAspArgProSerSerAlaGlySerThrProP 2462
               |||      |||                ::::||||||:::         :::   |||
Db    2805571 CACCGGCACCACCGGCGCCGCCGTTGCCGAACAGGCCCGCGTTGCCACCCGCCCCGCCGG 2805630

Qy       2462 hePro------------------------------------------TyrAsnProLeuI 2468
                |||                                             :::|||   :
Db    2805631 CGCCACCACCAGCCCCACCCATACCGCCGACGCCGCCCCCACCCGCCAGCCATCCACCCG 2805690

Qy       2468 leMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlyS 2488
```

```
              ::      :::      |||      |||   |||   ||||||      |||      :
Db    2805691 TCCCGCCGGTACCTCCGGGTGCGCCCGCCCCGCCCGCGCCACCGGCGCCGCCGATCCCAA 2805750

Qy       2488 erGlyProLeuAlaGlyPro 2494
              ::   |||   |||   |||
Db    2805751 ATAACCCGGCCGCCCCGCCG 2805770

RESULT 14
US-08-804-227C-1/c
; Sequence 1, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  43280 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  816..14234
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14351..19945
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20010..31199
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31232..36067
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36249..41774
US-08-804-227C-1

Alignment Scores:
Pred. No.:               1.25e-20       Length:        43280
Score:                   661.50         Matches:       696
Percent Similarity:      28.75%         Conservative:  275
Best Local Similarity:   20.61%         Mismatches:    1111
Query Match:             5.01%          Indels:        1310
DB:                      2              Gaps:          138

US-09-522-753-5 (1-2517) x US-08-804-227C-1 (1-43280)
```

```
Qy       54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
            ||||||   :::      |||   |||:::         ::::::|||
Db    12977 AGCCCGGCCGCCGCGGCTCAGCAGCAGCAGATGGCGCACACCGTGCTCGGCGGCCAGATG 12918

Qy       74 ProGly------------------------------------AsnGluArgSerGlnGlu 81
            ||||||                                       :::|||::::::
Db    12917 CCGGGCCAGCAGCCGCCCGAGCGTGCCGGTCGCGCCCGTCACCAGGACCGTGCCCGCCGG 12858

Qy       82 LeuHisLeuArg-ProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGl 101
            :::      ||| |||         |||   :::|||         ::::::   :::
Db    12857 GTCGAATGCCGGACCCCGCTGCTCCTCCGGGGTCCCGGCCCGGAGGAGGACCGGGGTCCG 12798

Qy      101 uPheIleGluSerLysArgProArgLeuGluLeuLeuProAspPro-------------- 116
                :::   ::::::|||         |||      |||   |||
Db    12797 GGCGGTGCCGGCCCGCAGG---------GAGAGCTGGCCGGCGCCGTCGAGGACCGCGCG 12747

Qy      117 -LeuLeuArgProSerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuTh 136
                   |||   |||   :::   ::::::|||   |||:::   |||
Db    12746 CAGCGCGAGGGGCAGCGCGGTGCGGTCGTCCGGATGCCCGTCGGTGTCGACGAGGGCGAA 12687

Qy      136 rLys-AspArgSerLeuThrGlyLysLeuGluProValSerProProSerProProHisT 156
             :::       :::   |||   :::      |||         |||:::|||||||||
Db    12686 GCGGCCCGGGTGCTCGGACCGCGCGGACCTCACCA------GGCCCCACACCCCCGCACA 12633

Qy      156 hrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnM 176
            ||                     |||   |||
Db    12632 CGCCAGGTCCCCGATCTCCTCCCCCGGTACGG---------------------------- 12601

Qy      176 etAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysL 196
                                                                 :::::::
Db    12600 -----------------------------------------------------CGGCGAC 12594

Qy      196 ysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSerP 216
            ::                     |||      ||||||   |||   :::|||   ||||
Db    12593 GGCTCCCCTGGTGAGCACTATGAGCGGTGTGTCCACCAACTCCTCGGCGTCCAGCCAGTC 12534

Qy      216 roProProIle-GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArg 235
            ||          ::::::      |||   |||
Db    12533 CTGGAGCGCGGTCAGACAGCGCTCCGTTGTCTG--------------------------- 12501

Qy      236 LysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuPro 255
                              ||||||:::      |||   |||
Db    12500 ---------------CCGCACCGACTCCGGAGTGGCCCCGGC------------------ 12474

Qy      256 LeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAla 275
                        |||   :::   |||||||||   :::      :::      :::|||
Db    12473 ------------CCCGGAGAGCACCGGCAGTAC---GACCGCTTCCGGCCGCGGCGCGCC 12429

Qy      276 MetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLys 295
               |||::::::|||               :::|||   |||   |||   |||
Db    12428 CGCCGCCAGCGCCTT------------CCGCAGCGCCCGCACCTCCGGGTACAGCGCGGT 12381

Qy      296 GlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArg 315
                        :::         |||               |||
Db    12380 GGGCCCACCGTCCGGCTGATCCCCCAGGCCCGGCGTCCCGGCGAGCCGGTCCGCGTCGGG 12321

Qy      316 IleGluAsnAsnPro---ArgArgArgAlaLysGluSerLysValArg---------Glu 331
               |||:::   |||      |||||||||         :::   |||         :::
Db    12320 GCCGAGCACCGCCCAGGAGCGCGGCGGGCCGGCGGGCAGGAGCCGCGGCTGCCACTCGAC 12261

Qy      332 TyrTyrGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGln 351
                  ::::::   |||      :::|||   ::::::   :::|||         :::
Db    12260 GCGGTACAGCGACCGTTCGCCGGCGTCCGTGCCGAGCAGCACGTCCAGCTTCCGGCGCGC 12201

Qy      352 SerArgValGlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGlu 371
               |||         :::|||                        :::   ||||||
Db    12200 CAGCGGACGCATCACCAGGGC---------------------CCCCAAGGAGAGCACGGC 12162

Qy      372 ValSerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGln 391
               :::      :::   ||||||
Db    12161 CGCGCCGGACGTGTCTGCGGCCTGGAG--------------------------------- 12135
```

```
Qy      392 LeuAlaValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsn 411

Db    12135 ------------------------------------------------------------ 12135

Qy      412 MetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMet 431
                                                :::||||||   |||      :::
Db    12134 ------------------------------------CGAGACCGCCCCGTCACCCGCCTT 12111

Qy      432 TrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPhe 451
                           :::   |||         :::      :::|||
Db    12110 CGTCAGCCGCACCCGCAGAGCACCGGCACCCGTCGCGTAGAGCGACACTCC--------- 12060

Qy      452 GlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyr 471
                              |||      :::            |||||||||   ::::::
Db    12059 ------------------CTGCCAGGCGAAGGGCAGCAGCACTGCGTCCTCCGGCACTTC 12018

Qy      472 LeuThrLysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyrArgArg------ 489
                  ::::::                  |||   ||||||      :::|||
Db    12017 ------GAGGAGACCACCGGCGATCCACGGCTGCAGCGCCGCGTCCAGCAGCGCGGGTG 11964

Qy      490 ------------------------ArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGln 501
                                    ||||||         |||:::         :::|||
Db    11963 CACACCGAACCGGCCGGCACCGGTCGAGGCCCGGTCGGGCAGGAGCACCTCGGCGAGCAG 11904

Qy      502 ---------------------------------GlnGlnGlnGlnGlnGlnGlnGlnGln 510
                                             :::::::::      :::   :::
Db    11903 CGCGTCCGGCTGCCGCCATACGGCGCGCACGCCCGCGAACGCCT-CACCGTACTCGTAGC 11845

Qy      511 ProMetProArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGlu 530
            |||   ||||||:::   :::      :::      :::         :::   |||
Db    11844 CCAGGGCCCCGAACCGTTCGTAGAGCTCGGCCGGTTCGACGCGTTCCCCGGCGGGCGGCC 11785

Qy      531 Lys------GluGluGluLysProGluVal------------------------------ 538
            :::               :::||||||
Db    11784 AGGCCCGCGCCCAGTCGGCGTCCGGAGGCGTGGGCTCCGGGTCCGGCACGAGCGTGCCCT 11725

Qy      539 ------------------GluAsnAspLysGluAspLeuLeuLysGluLysThrAspAsp 552
                                                ||||||:::
Db    11724 CGCCGTGCCGCGTCCAGACACCGGCCGAGTCGCCGCCTTCTGCGGCCGAGTGGACGGCGG 11665

Qy      553 ThrSerGlyGluAspAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThrAla 572
                  |||                     ||||||   ::::::   |||   |||
Db    11664 CTTCCGGGGCCGAGTGGATGGCGACCCTCCGCCGTCCCGACTCGTCGGCGGCCGACACCC 11605

Qy      573 AsnSer---------GlnGlyArgArgLysGlyArgIleThr------------------ 583
            :::|||               ||||||:::   |||   |||
Db    11604 TCATCTGCACCCGCACACCACCGGCGGCGGGCAAGACCAACGGCGCCTCAACGCTCAGCT 11545

Qy      584 ------------------------ArgSerMetAlaAsnGluAlaAsnSerGluGluAla 595
                                    |||      |||      |||||||||   :::|||
Db    11544 CTTCCAGCCGCCCGCAGCCGGTGGCGGCT-CCCGCGTGCAGCGCGAACTCCACGAAGGCC 11486

Qy      596 ---------------IleThrPro---------------------GlnGlnSerAlaGlu 603
                           :::||||||                     |||      :::|||
Db    11485 GTACCCGGCAGCAACGTCACCCCGGACACGGCATGGTCCGTCAGCCAGTCGTGCTCGGAG 11426

Qy      604 LeuAlaSerMetGluLeuAsn------------GluSerSerArg--------------- 614
               ||||||         |||            ||||||:::|||
Db    11425 CCCGCCAGCCGCCCGGCGAACACCCGCTCACCGGAGTCGGCCAGGGGTAACCCGCCACGG 11366

Qy      615 ---------TrpThrGluGlu--------------------------------------- 618
                     |||:::   :::
Db    11365 AGGAAGGGGTGGTCCTCCCAGGTCATACCGAACCGGCCCGCAGCCGCGTCGCCGACTCCG 11306

Qy      618 ------------------------------------------------------------ 618

Db    11305 GCGGGGGATGTCACGGCCCAGTAGTGGTGGTGGTTGAAGGGGTAGGTGGGGAGGGTGGTG 11246

Qy      618 ------------------------------------------------------------ 618

Db    11245 AGGTGACCTGGGGGCAGGGCGGGGTCCCAGGTGAGGGTGGCGCCGTGGGCCCAGGCCTCG 11186

Qy      619 GluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAla 638
```

```
                 :::|||   :::::::::      |||         |||            :::|||
Db        11185 GCGGTGGAGGTCAGCAGGCGGTGCGGGCTGTCGTCCTGGCGGCGCAGGGTTCCGGTGGCA 11126

Qy          639 ArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLys 658
                      :::   |||         :::   |||   |||            |||:::
Db        11125 TGGACACTCCGGTCCGCGGCCTCGGCCGTCTGCTC-AATGGCATGGACGAGTACCGGGTG 11067

Qy          659 ArgGlnAsnLeuAspGlu------------IleLeuGlnGln-HisLysLeuLysMetGl 674
                      :::|||||||||            :::|||||||||                ||
Db        11066 GGCACTGACCTCGATGAACCGACGGTGGCCCTGCTCCAACAAACCACCGACCACCGCGGA 11007

Qy          674 uLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAl 694
                |               ||||||   :::   :::   ||||||:::   :::
Db        11006 GAACTCAACCCGGTTCCGCAGATTACGGAACCAATACCCCGCATCGAAAACCGGCTCGCC 10947

Qy          694 aAlaPheProProValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGl 714
                          |||            |||               |||:::   |||
Db        10946 CGGCTGCTCTCCGGCGACGGTG---GAACACACCGGCACCCGCGGACTGCGGGGGCGGAT 10890

Qy          714 uGluGluMetValGluGluAlaGluAlaLeuHisAlaSer-------------------- 727
                    :::                           :::|||
Db        10889 ACCGGACAGCACCCGCTCCAACTCGCCCTTCAGGTCCTCCACATGGCGGGAGTGAGAGGC 10830

Qy          728 ----GlyAsnGluValProArgGlyGluCysSerGly----------------------- 738
                             :::   |||      |||:::
Db        10829 GTAGTCCACCGGTATCAACCGAGCC---TGCACACCCTCCGCCTGACAGACGGCCACATA 10773

Qy          739 -ProAlaThrValAsnAsnSerSerAspThr----------------------------- 748
                 |||||||||         |||   :::|||
Db        10772 ACCGGCCACCGCCCGCCGATCCCCGGAAACCACCGTGGACGCCGGACCGTTGACCGCGGC 10713

Qy          749 ----------------------GluSer-IleProSerProHisThrGluAlaAlaLys- 760
                                         |||    |||   |||      :::||||||:::
Db        10712 TACCTGTACTCCGGGCCACTTCGCCAGCCCGGCCTCGACCTCGCCGGCAGGCAGCGGAAC 10653

Qy          761 --------------------------AspThrGlyGlnAsn------------------- 765
                                          ||||||||||||:::
Db        10652 AGCCGCCATCGCACCACGACCCGCCAGATACCGGCCAATCAACCCCGCCCGCAGAGCGAC 10593

Qy          766 -----GlyProLysProProAlaThrLeu------------------------------- 773
                        |||   ||||||||||||
Db        10592 CACAGCCGCCGCATCCTCCAGCGACAACGCCCCCGCCACCGTGGCAGCAGCGATCTCCCC 10533

Qy          774 --GlyAlaAspGlyProPro------------------------------ProGlyProP 783
                  |||   |||   ||||||                                 ||||||
Db        10532 CTGGGAATGACCCACCACCGCAGCCACGTCCACACCCATCGCCTGCCAGTACCGGGCCAG 10473

Qy          783 roThrProProArgArgThrSerArgAla------------------------------- 792
                  ||||||         ||||||||||||
Db        10472 CGACACCATCACCGCCCACGTCACGGGCTGCACCACATCCACCCGCTCCAACCCGGCCCC 10413

Qy          793 --------------ProIleGluProThrPro---------------------------- 798
                              |||      |||||||||
Db        10412 GCCCGACACCACCTCCAGCAAATCCCACTCCACATGAACACTCAGCGCCCGCGCACACTC 10353

Qy          798 ------------------------------------------------------------ 798

Db        10352 CCGCATCGACGCCGCAAACACCTCCGAAGACCCCAGCAGCCCCGCACCCATCCCCACCCA 10293

Qy          799 --AlaSerGluAlaThrGlyAlaProThrPro---------------ProProAlaPro- 812
                  ||||||         |||   |||||||||                  |||:::|||
Db        10292 CTGCGTCCCCTGACCAGGGAAAACCAACACCACGCCCCCACCACCGGACCCCGTCTCCGG 10233

Qy          813 --------------------------ProSerProSerAlaProProProVal------- 821
                                          |||   |||      |||||||||
Db        10232 GTCCACCACCGCACCGGTGACCACACCCGCGTCCGGATCACCCCCCGCCAACGCCCGCAA 10173

Qy          822 --------------------------ValProLysGluGluLysGluGluGluThrAlaA 833
                                             |||   :::      :::   :::||||||
Db        10172 ACGCCCGGCAAGCTCAGCCTCACCCCGAGCCAGCACAACCGCCCGATACCCAAACCGCGA 10113

Qy          833 laAlaProProValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaV 853
                     ||||||                        |||||||||            |||
Db        10112 CCGGCCCGCCA---------TCGTCACCGCCACATCCGCCAGCCCCACACCACCG-GCC- 10064

Qy          853 alAspThrGlyLysAlaGluGluProValLysSerGluCysThr------GluGluAlaG 871
                     |||      |||      |||      |||      |||      |||      |
Db        10063 --TCCACCACCTCCGCCAACCGGCCCGCAGCCTCCCGCACCACCACCCGCGAACGCCCCG 10006

Qy          871 luGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaL 891
                ||                        |||      |||   :::         ||||||
Db        10005 AAACCACCACCGGCACCACACCCACAGCAGGCTGTGCCTCTTCGGGACTGCCAGGAGCCT 9946

Qy          891 eu------LysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerS 909
                           |||            |||   |||      :::   |||:::   :::|
Db         9945 TCGCACCCTCCGCCTCGGTGGCGTCGGGGGCTTCGAGGGCTTCGGTGACGTCGGTGACGT 9886

Qy          909 erGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluValAspGluA 929
                ||||||||         |||   :::|||            |||               |
Db         9885 CGGGGGCTTCGGGGGCTTCGGGGGCTTCGGGGGCGTCGGTGGCTTCGAGGGCTTCGAGGG 9826

Qy          929 laGluGlyGlyAspLysAsnArg---LeuLeuSerProArgProSerLeuLeuThrProT 948
                ||   |||      :::   |||      |||   |||            |||||||||
Db         9825 CTTCGGGGGCTTCGAGGATCAGATGGCGTTGGTACCA-------------CTGACACCGA 9778

Qy          948 hrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLysG 968
                  |||      |||            |||      |||      |||   :::|||   |
Db         9777 A-GGCGGAGACCGCCGCCCGACGCGGCCGGTCCACCCCCGGCCACTCCCGCGCCTCGGTC 9719

Qy          968 lnArg---------------------------------AlaAlaAlaIle---------- 973
                ||:::                                 |||      |||
Db         9718 AGCAACCGCACACCACCAGCCGACCAGTCCACATGCCGGCTCGGCTCATCCACATGCAAC 9659

Qy          974 -----------ProProIleGlnValThrLysVal------------HisGluProPro- 985
                           |||||||||      :::                  |||:::||||||
Db         9658 GTCCGCGGCAACCACCCATACC-GCATCGCCAACACCATCTTGATCACACCAGCCACCCC 9600

Qy          985 ------------------------------------------------------------ 985

Db         9599 CGCAGCCGCCTGAGCATGACCAATATTCGACTTCAACGACCCCAACCACACCGGACACTC 9540

Qy          986 --ArgGluAspAlaAlaProThrLysProAlaProProAlaPro---------------- 999
                        |||   |||   |||   ||||||   |||||||||
Db         9539 AGCATCACGATCCCGCCCATACGTCCCCAGCAACGCCTGCGCCTCGATCGGATCACCCAG 9480

Qy         1000 --------------------ProProProGlnAsnLeuGlnProGluSerAspAlaProG 1013
                                    |||||||||            |||   |||   |||
Db         9479 CCGCGTCCCCGTCCCATGCGCCTCCACCACATCCACATCCGCCACCGTCAACCGCG---- 9424

Qy         1013 lnGlnProGlySerSerProArgGlyLysSerArgSerProAlaProProAlaAspLysG 1033
                              ||||||               ||||||||||||      |||
Db         9423 -----------CATTCGCCAACG---CCTGACGAATCACCCGCTCCTGCGACGGACCACT 9378

Qy         1033 luAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyL 1053
                  |||   :::      :::         |||      ||||||      |||
Db         9377 CGGCGCCGTCAAACCATTCGACGCACCGTCCTGATTCACCGCCGACCCACACA------- 9325

Qy         1053 euProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspProS 1073
                  |||   |||   |||                     :::|||   |||      |||:
Db         9324 --CCACCGCCAACACCGGAT------------------GCCCCAACCGCCGGGCATCCGA 9285

Qy         1073 erAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaA 1093
                ::            |||||||||   ||||||   |||         |||
Db         9284 CAACCGCTCCACCAGCACCACCCCCGCACCCTCGGCCCAAGCGGTCCCAT---------- 9235

Qy         1093 rgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAlaL 1113
                  |||      ||||||   |||      ::::::   ||||||      ::::::||||
Db         9234 --CCGCCCATCCGCGAACGCCTTGCACCGCCCATCGGCCGCCAACCCCCGCTGCCGCGA 9177

Qy         1113 ysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnL 1133
                ||   |||                           |||            |||
Db         9176 AAACTCCACGA---------------------ACAATCCGGGTGCCGCCATCACCGTCAC 9138

Qy         1133 euHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuP 1153
                  |||   |||         |||   :::   |||      |||
Db         9137 ACCACCGGCCAACGCCATGTCACACTCGCCACCCCGCAACGCCTGCG------------- 9091
```

```
Qy     1153 roLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuS 1173
                                :::   |||                        :::
Db     9090 -----------------CCGCGTTGTGCAGGGCGACCAGCGACGACGAACACGCGGTGTC 9048

Qy     1173 erProArgGlyGlnAlaGlyPro-ProGluSerLeuGlyValProThrAlaGlnGluAla 1192
                 |||   :::      ||| |||:::   :::                  :::|||
Db     9047 GATGGAGAGCGAAGGACCCTCCAGCCCCAGCGTGTACGAGATGCGGCCGGAGACGACGCT 8988

Qy     1193 SerValLeuArgGlyThrAlaLeuGlySerVal---------------------ProGly 1205
                  |||   |||      :::      |||                     ||||||
Db     8987 CGCCGACTTGCCGGTGAGCAGATAGCCTTCGTACTCCTCCGGGGCCTCGTGCAGCCGGGA 8928

Qy     1206 GlySerIleThrLysGlyIleProSerThrArgValProSerAspSerAlaIleThrTyr 1225
            |||   :::      |||   |||                                 |||
Db     8927 GGCGTACTCCTGATGGGTCAGCCCGAC---------------------------GTATAC 8895

Qy     1226 ArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArg 1245
               |||:::                                          :::   |||
Db     8894 CCCGGTGCG------------------------------------GCTGCCGTGCAGCGA 8871

Qy     1246 IleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLys 1265
            :::                     |||:::|||            |||   ||||||
Db     8870 GTG---------------------CGGGTCGATACCCGCCCGCTCGATCGCCTCCCAGGA 8832

Qy     1266 Gly---------------------HisVal-IleTyrGluGlyLysLysGlyHisValLe 1278
            |||                        ::: |||            :::|||   :::
Db     8831 GGTCTCCAGCAACAGCCGCTGCTGCGGATCCATCGCCACCGCCTCACGCGGCGAAATCCC 8772

Qy     1278 uSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSer-SerS 1298
                               |||               |||         |||:::    |
Db     8771 GAAGAACTCCGCGTCGAACTCCGCCGCGTCGTGGAGAAAACCCCCCTCCCGCACATATGT 8712

Qy     1298 erGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgV 1318
            ||   |||||||||               :::|||                  ||||||
Db     8711 CTTCCCCGCCACACCCGGCTCCGGGTCGTAGACGGAATCCAGGTCCCAGCCCCGGTCGGT 8652

Qy     1318 alGlyArgAlaIle------------------SerSerAlaSerIle------------- 1327
                 |||   :::                  |||   ||||||
Db     8651 GGGCGCGGCGCCTATGACGTCACGCTCCTCGGTCACCAGCTTCCACAGCGCCTCAGGGGT 8592

Qy     1328 -----------------GluGlyLeuMetGlyArgAlaIleProProGluArgHisSerP 1342
                                |||         |||      ||||||   |||      |
Db     8591 GCCCGCCCCACCGGGGAAACGGCACGCCATGCCGACGATGACCACCGGGTCGTCGTCGGC 8532

Qy     1342 roHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgS 1362
            ||
Db     8531 CACGG------------------------------------------------------- 8527

Qy     1362 erTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArgGluG 1382
                                         ||||||               :::|||   |
Db     8526 -----------------------------CGACGGGCAGTGGCGGCTCATCGGCGGCG-G 8497

Qy     1382 lyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaL 1402
            ||         |||   |||   ||||||      :::         :::         |
Db     8496 GCGCGGGGGCTCCGGCGCCGGTCAGCCGGTCGAGCAGCAGTCTGGCCAGGGCGGTCGGGC 8437

Qy     1402 euGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaG 1422
            |||||      |||:::                           ||||||      :::
Db     8436 TGGGGTGGTCGAAGATG---------------------------ACGGTGGCGGGTAGCT 8404

Qy     1422 lyArgSerIleHisGluIleProArg------------GluGluLeuArgHisThrProG 1438
              |||   :::      :::   |||               |||:::|||   :::
Db     8403 TCAGGCCTGTGCCGGCATTGAGGCGGTTGCGGAGTTCCAGGGAGGTCAGGGAGTCGAAAC 8344

Qy     1438 luLeuProLeuAlaProArgProLeuLysGluGlySerIle---------------ThrG 1453
                    |||      |||               |||:::               |||:
Db     8343 CGGCCGACTTGAACGGGCGTTCGGCGCCGATGGCGTCCGTGTTCGGATGCCGCAGCACAA 8284

Qy     1453 lnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspV 1473
            ::         ::::::               |||:::|||      ::::::      |
Db     8283 AGGCCGCCTCGGTACGG------------ACGAGTTCGACCAGCTTCCGCAGGCGTTCGG 8236
```

```
Qy     1473 al---------------------ArgSerLeuIleGlySerProGlyArgThrPheProP 1486
            ||                     |||:::      |||:::|||
Db     8235 TCTCGTTCAGCATGGACAACCGCCGCGCCAGCTCGGGCGCGCCGTCCGGCGTC------- 8183

Qy     1486 roValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluG 1506
                          ||||||:::      |||            ||||||
Db     8182 -----------GCGGACGTGGTGCCGGCTGCCGTCAGGGCCTCCCGCGCCTCCGGCACGG 8134

Qy     1506 luSerLeuLysSerArgProGly------------------------------------- 1513
              |||      ||||||||||||
Db     8133 TGTCG---AACAGCCGCCCGGGACGGATGGCGTTCGTCGCGGGCGCGAAACGCTCCCAGT 8077

Qy     1514 -----------------------------------------------------ThrAlaS 1516
                                                                    ::::
Db     8076 CGACGTCGACGACAGTGACGCACACGTCGTTCCTGCCCATGGCGCCCAGGAGCGCGTCGA 8017

Qy     1516 erSerSerGlyGlySerIleAlaArgGly---------------------------AlaP 1527
            ::::::::   ||||||:::||||||                              ||||
Db     8016 CGGCCGCGTCGGGGTCCATGGCCCGCAGCCCTCGCCGCGACAGACTCTCCTCGCCCGCAC 7957

Qy     1527 roValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisG 1547
            ||      :::|||            |||   |||   |||         |||      |
Db     7956 CCGCCGCCATGCCT---------CCCCCGCCCCAAAGGCCCCAGGCGACGGAGGTCGCGG 7906

Qy     1547 lyAlaProPheAla---------------------------------------------- 1551
            ||:::|||   |||
Db     7905 GCAGTCCGGCGGCACGGCGACGCTCGGCGAGGGCGTCGAGCGCGGCGTTGGCGGCGGCGT 7846

Qy     1552 --------------------GlyHisLeuProArgGlySerProValThrMetArgGluP 1565
                                   |||:::|||                     :::
Db     7845 ACGCACCCTGGCCGGCGTTGCCCCATGTGCCGGTGACGGAGGAGAAGAGGACGAAGGCGT 7786

Qy     1565 roThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysL 1585
                 |||   :::         |||      ||||||   |||
Db     7785 CCAACCCTTTTATGTCCGCAGTCAGTTGGTGCAGCAGCTCCGCGCCGCAC---------- 7736

Qy     1585 euThrSerThrProArgGluIleAlaLys-SerPro------------------------ 1596
              |||      ||||||   ::::::||| ||||||
Db     7735 --ACCTTCGCCCCGCGGACGGTCTCGAAGCTCTCCGGTGACAGCGTGTCGATCACCGCGT 7678

Qy     1596 ------------------------------------------------------------ 1596

Db     7677 CGTCCAGAATCCCGGCGGTGTGGAAGACGGCGTTCGGCGGATACGCGGTGACGAGCGCGG 7618

Qy     1597 ------HisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeu 1614
                  |||         |||   |||   |||            |||      |||:::
Db     7617 CGAGCGCATCACGCTCGGCCACGTCACAGGCCGCGT---GCACGACCTCGCAGCCATGTC 7561

Qy     1615 Leu--------------------------------------------------------- 1615
            |||
Db     7560 CTCGCAGTTCCTCGGCGAGTTCGGCGGCGCCCGGCGCCTCCGGTCCGCGACGGCTGGTGA 7501

Qy     1616 ------------ArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
                        |||      |||         |||      |||   :::|||:::|||
Db     7500 GGACCAGGCGTTCGGCT-CCCTCGGCCGCCAGCCTGCGGGCGAGCCGCCTGCCGATGGCA 7442

Qy     1632 PheAspPro---------ThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAla 1648
                  |||         :::::::::|||   |||         :::||||||   |||
Db     7441 CCCATGCCGCCGGTGATCAGCACCGTGCCGGACGGCTGCCAG---GAGGCGGCCCCCGCT 7385

Qy     1649 TyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyr 1668
                     |||||||||      |||   |||      |||
Db     7384 GCCGCGGC-CCGCGACAC------CCGCCGCCCGTAGATCCC------------------ 7350

Qy     1669 LeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAsp 1688
                  ||||||
Db     7349 ------CGAGGC------------------------------------------------ 7344

Qy     1689 TyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAsp 1708
                                 |||   :::         |||:::|||         |||
Db     7343 ------------------GCGCACGGCGACCTGGTCCTCGCCGTCGCGGCCGGCGAGGAC 7302

Qy     1709 ---------MetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAla 1725
```

```
                           |||:::|||   |||      ||||||:::   |||         :::
Db        7301 GGCGGCAGCAAGCTCAAGGGCGCGTCGGT-TCGGAGTCTGCGGCAGCCTCTCCGGGGTCT 7243

Qy        1726 AlaGly------------------------------ProArgGlyIleIleAspLeuSer 1735
                  |||                              |||            |||   |||
Db        7242 GTGGGGGCTGTGAGGTCTCCGGCGTCCGCGTCAGCTCCGCTGTCTCCGGCAGATCGATCA 7183

Qy        1736 GlnValProHisLeuPro---ValLeuValProProThrProGlyThrProAlaThrAla 1754
                     |||      |||   |||      ||||||         :::|||   |||
Db        7182 AGCCGCCCCAGATATCCGGGAGTTCAAGGGCCGCCACTCGTCCGAAGCCCCACAACTGGG 7123

Qy        1755 MetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSer 1774
                     |||   |||      |||      |||:::|||      |||      :::
Db        7122 CCCCCGCGTCCTGCGGCGTCTCCCCGGGAACCACGGCCACGGCCCTCCGCGTCAGCGCCC 7063

Qy        1775 SerProLeuSerProGly------------------------------------------ 1780
                  |||      ||||||
Db        7062 AGACCCTGGCGGCCGGGACGCCGATGTCGGAGACCGCCTGGAGGAGCGCTATCGACGTCC 7003

Qy        1781 ---------------------------------------------------GlyProThr 1783
                                                                     |||
Db        7002 TGACCGCGTCCGTGGAGGGCACCAGACCGAGAAGCGAGACGACGGTGGTGTTCTCCTCCC 6943

Qy        1784 HisLeuThr--------------------------------------------------- 1786
                  ||||||
Db        6942 CCGCTGACTCCATGGAGGGCTTGAGCAGTTCGGCGACCTCGGCACGCCGCGCGCCGATGG 6883

Qy        1787 ---------------------------LysProThrThrThrSerSerSerGluArgGlu 1797
                                          :::      ||||||      |||   |||
Db        6882 GCACCTCAAGGCGGGCGACGGAGGCACCGCGCTCGGACAACACCGTTGTCAGTCCGTCGA 6823

Qy        1798 ArgAspArg---------------------------------AspArgGluArgAspArg 1806
                     |||                                 ||||||      :::
Db        6822 GGAGCTCGCCGTCGGTGGTCGGGGGTGCAACGAGGAGGCAGCGACCGATGGAGTGAGGGC 6763

Qy        1807 AspArgGluArgGluLysSerIleLeuThrSer---ThrThrThrValGluHisAlaPro 1825
                  |||   |||   :::|||   |||:::            :::      |||
Db        6762 GGACGGGGGAGGATTCGGTCAGGGCTTTCCAGGTGACGCGGTAGCGCCAGGCATCGG--- 6706

Qy        1826 IleTrp------------------------------------------------------ 1827
                  |||
Db        6705 ---TGGGGCTCTGGGTGGTCGTCGCGGGGGTGGTGGGGGTGGTGTCGAGCCAGTAGTGGT 6649

Qy        1827 ------------------------------------------------------------ 1827

Db        6648 GGTGGTTGAAGGGGTAGGTGGGGAGGGTGGTGAGGTGGCCTGGGGGCAGGCGGGGTCCC 6589

Qy        1828 ------------------------ArgProGlyThrGluGlnSerSerGlySerSerGly 1839
                                       ||||||         :::|||:::|||   :::|||
Db        6588 AGGTGAGGGTGGCGCCGTGGGCCCAGGCCTCGGCGGTGGAGGTCAGCAGGCGGTGCGGGC 6529

Qy        1840 SerSerGlyGlyGlyGlyGlySer------------------------SerSerArgPro 1851
                        ||||||   ||||||                              ||||||
Db        6528 TGTCGTCCTGGCGGCGCAGGGTCCCGGTGGCATGGACACTCCGGTCCGCGGCCTCGGCCG 6469

Qy        1852 AlaSerHisSerHisAla------------------------------------------ 1857
               ::::::   ||||||
Db        6468 TCTGCTCGATCGCATGGACGAGTACCGGGTGGGCACTGACCTCGATGAACCGACGGTGGC 6409

Qy        1858 ---------------HisGln-His---------------------------------Se 1861
                              |||::: |||
Db        6408 CCTCCTCCAACAAACCACCGACCACCGCGGAGAACTCAACCCGGTTCCGCAGATTACGGA 6349

Qy        1861 rProIle----------------------------------------------------- 1863
                ||||||
Db        6348 ACCAATACCCCGCATCGAAAACCGGCTCGCCCGGCTGCTCTCCGGCGACGGTGGAACACA 6289

Qy        1864 ----SerProArgThr------------------GlnAspAlaLeuGlnGlnArgProSe 1876
                      |||||||||                  |||      ||||||
Db        6288 CCGGCACCCGCGGACTGCGGGGGCGGATACCGGACAGCACCCGCTCCAACTCGCCCTTCA 6229

Qy        1876 rValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer-------- 1893
                |||||||||      |||::::::|||::::::         :::||||||
```

```
Db        6228 GGTCCTCCACATGGCGGGAGTGAGAGGCGTAGTCCACCGGTATCAACCGAGCCTGCACAC 6169

Qy        1894 ----------------------------------LysProThrValLeuArgSerThr-- 1901
                                                 :::|||   :::   |||
Db        6168 CCTCCGCCTGACAGACGGCCACATAACCGGCCACCGCCCGCCGATCCC-CGGAAACCACC 6110

Qy        1902 -------------------------------------SerThrSerSerProValArgPr 1909
                                                          :::||||||   |||||
Db        6109 GTGGACGCCGGACCGTTGACCGCCGCGACCTCCACACCCGGCCACTTCGCCAGCCCGGCC 6050

Qy        1909 o---------------AlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyTh 1924
               |               ||||||      ||||||:::   |||   |||         ||
Db        6049 TCGACCTCGCCGGCAGGCAGCGGAACAGCCGCCATCGCACCACGACCCG---CCAGATAC 5993

Qy        1924 rLeuAspGlyValTyrPro---------------------------------ThrLeuMe 1933
               |               |||                                 |||
Db        5992 CGGCCAATCAACCCCGCCCGCAGAGCGACCACAGCCGCCGCATCCTCCAGCGACAACGCC 5933

Qy        1933 tGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArg-- 1952
                   |||            :::      |||            |||   :::||||||
Db        5932 CCCGCCACCGTGGCAGCAGCGATCTCCCCCTGGGAATGACCCACCACCGCAGCCACGTCC 5873

Qy        1953 -------------AlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGl 1968
                            |||   ||||||   |||      :::   |||   :::|||:::
Db        5872 ACACCCATCGCCTGCCAGTACCGGGCCAGCGACACCATCACCGCCCACGTCACGGGCT-- 5815

Qy        1968 yLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProProVa 1988
                         |||      |||         :::   |||||||||      ||||||
Db        5814 ----------GCACCACATCCACCCGCTCCAACCCGGCCCCGCCCGACACCACCTCCAGC 5765

Qy        1988 lSer------------GlyHisAlaThrIleAlaArgThrProAlaLysAsn-------- 2001
                :::               |||:::      |||   |||||||||
Db        5764 AAATCCCACTCCACATGAACACTCAGCGCCCGCGCACACTCCCGCATCGACGCCGCAAAC 5705

Qy        2001 ------------------------------------------------------------ 2001

Db        5704 ACCTCCGAAGACCCCAGCAGCCCCGCACCCATCCCCACCCACTGCGTCCCCTGACCAGGG 5645

Qy        2001 ------------------------------------------------------------ 2001

Db        5644 AAAACCAACACCACGCCCCCACCACCGGACCCCGTCTCCGGGTCCACCACCGCACCGGTG 5585

Qy        2001 ------------------------------------------------------------ 2001

Db        5584 ACCACACCCGCGTCCGGATCACCCCCCGCCAACGCCCGCAAACGCCCGGCAAGCTCAGCC 5525

Qy        2001 ------------------------------------------------------------ 2001

Db        5524 TCACCCCGAGCCAGCACAACCGCCCGATACCCAAACCGCGACCGGCCCGCCATCGTCACC 5465

Qy        2002 ----------LeuAlaProHisHisAlaSerProAspProPro----------------- 2012
                            ||||||||||||      |||   ||||||
Db        5464 GCCACATCCGCCAGCCCCACACCACCGGCCTCCACCACCTCCGCCAACCGGCCCGCAGCC 5405

Qy        2013 -------AlaProProAlaSerAlaSerAspPro------------------------Hi 2022
                         |||||||||:::|||      |||                        ||
Db        5404 TCCCGCACCACCACCCGCGAACGCCCCGAAACCACCACCGGCACCACACCACTGCGGGCA 5345

Qy        2022 sArg------GluLysThr---------GlnSerLysProPheSerIleGlnGluLeuGl 2037
               ||||      ::::::|||         ||||||   |||   |||   :::
Db        5344 TCGAGGACCGCAGCAGACTCCGACACCTCAGAGCGGACCGGGGTCGTGGCGCTCTCCGCC 5285

Qy        2037 uLeuArgSerLeuGlySerTyrHisGlySerSer----------------------------- 2047
                   |||      |||      ||||||
Db        5284 TCGGCGGTGTCGGGGGCTTCGAGGATCAGATGGGCGTTGGTACCACTGACACCAAAGGCG 5225

Qy        2048 ----------TyrSerProGluGlyValGluProValSerProValSerSerProSerLe 2064
                                     |||      |||   :::|||            :::
Db        5224 GAGACCGCCGCCCGACGCGGCCGGTCCACCCCCGGCCACTCCCGCGCCTCGGTCAGCAAC 5165

Qy        2064 uThrHisAspLysGlyLeuProLysHisLeu-GluGluLeuAspLysSerHisLeuGluG 2084
                ||||||   :::            |||          |||         |||::::::
Db        5164 CACACACCACCAGCCGACCAGTCCACATGCCGGCTCGGCTC------ATCCACATGCAA- 5112
```

```
Qy          2084 lyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuP 2104
                         ||||||:::||||||            :::               ||||||
Db          5111 --------CGTCCGCGGCAACCA---------CCCATACCGCATCGCCAACACCATCTTG 5069

Qy          2104 roHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaP 2124
                   |||   :::|||   |||:::                              |||   |
Db          5068 ATCACAC-CAGCCACCCCCGCAGCCGCCTGAGCATGACCAATATTCGACTTCAACGACCC 5010

Qy          2124 roGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleT 2144
                 ||                                                         |
Db          5009 CA------------------------------------------------------ACCA 5004

Qy          2144 hrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuT 2164
                 ||   |||         |||   |||         |||         |||||||||
Db          5003 CACCGGACGCCCACCATCACGATCCCGCCCATACGTCCCCAGCAACGCCTGCGCCTCGAT 4944

Qy          2164 yrSerPheProGlyAlaSer---------CysProValLeuAspLeuArgArgProProS 2181
                         |||   ||||||            |||                  |||||||
Db          4943 CGGATCACCCAGCCGCGTCCCCGTCCCATGCGCCTCCACCACATCCACATCCGCCACCGT 4884

Qy          2181 erAspLeuTyrLeuProProPro------------------------AspHisGlyAlaP 2193
                 ||      :::   |||   |||                        ||||||   ||||
Db          4883 CAACCGCGCATTCCCCAACGCCTGACGAATCACCCGCTCCTGCGACGGACCACTCGGCGC 4824

Qy          2193 roAlaArgGlySerProHisSerGluGlyGly------LysArgSerProGluProAsnL 2211
                 ||:::      |||   |||                        :::|||   |||
Db          4823 CGTCAAACCATTCGACGCACCGTCCTGATTCACCGCCGACCCACACACCACCGCCAACAC 4764

Qy          2211 ysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGlu---- 2229
                      :::         |||                  |||      ||||||
Db          4763 CGGATGCCCCAACCGCCGGGCATCCGACAACCGCTCCACCAGCACCACCCCCGCACCCTC 4704

Qy          2230 -----GlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgA 2248
                            :::   |||   |||   |||   |||            |||:::|||
Db          4703 GGCCCAAGCGGTCCCATCCGCCCCATCCGCGAAC-GCC------------TTGCACCGCC 4657

Qy          2248 spGlyGluGlnThrGluProSerArgMetGlySerLysSer---------ProGlyAsnT 2265
                                        |||      |||   :::         |||      |
Db          4656 CATCGGCCGCCAACCCCCGCTGCCGCGAAAACTCCACGAACATCCCCGGCCCGGCCATCA 4597

Qy          2265 hrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysS 2285
                 ||      |||||||||                        |||:::   :::      :
Db          4596 CCGTCACACCACCGGCC------------------------AACGCCAT-GTCACACTCG 4562

Qy          2285 erLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArg---------------- 2299
                 ::      |||   :::::::::::::::      |||   |||
Db          4561 CCACCCCGCAACGCCTGCACCGCCAGATGCAACGCCACCAACGACGACGAACACGCCGTA 4502

Qy          2300 -----------------------------------------AsnGluProGluTyrAsnI 2306
                                                          :::   |||   ::::::
Db          4501 TCCACGGTCACGGCCGGACCCTCCAGCCCCAGCGTGTAGGCCACACGCCCCGACACCACA 4442

Qy          2306 leSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetT 2326
                      :::         |||         :::   |||            |||
Db          4441 CTCCCCGACGTACCGGTGAGGAGGTAGCCGTCGCTGCCTTCCGCTCCTTCGGC------G 4388

Qy          2326 hrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleI 2346
                      ||||||      ::::::   ||||||      :::   |||      |||:::
Db          4387 AGCCGAGGTCCGTATTCCTGCGGCATCACGCCGGCGTAGACGCCGGTGCGGCTGCCGTGC 4328

Qy          2346 leArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaA 2366
                   |||   |||      |||:::|||               |||||||||
Db          4327 AGCGAGT-GCG------GGTCGATACCCGCCC---GCTCGATCGCCTCCCAGGAGG---- 4282

Qy          2366 snAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaA 2386
                         :::|||      |||:::|||:::   |||:::      |||      |||
Db          4281 --------TCTCCAGCAACAGCCGCTGCTGCGGATCCATCGCCACCGCCTCACGCGGCGA 4230

Qy          2386 laAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValS 2406
                         |||:::   |||      |||   |||               |||:::
Db          4229 AATCCCGAAGAACTCCGCATCGA---ACTCCGCCG---------CATCATGCAGAAAACC 4182

Qy          2406 erGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspA 2426
                               :::               ||||||||||||      :::
Db          4181 CCCCTCCCGCACATACGACTTCCCCGCCACACCCGGCTCCGGGTCGTACAACC------- 4129

Qy          2426 rgProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuT 2446
                   ||||||      |||      |||   |||            |||      |||
Db          4128 --CCTCCACATCCCAGCCACGATCACCGGGAAACGGGAAACCGCGTCCACCCCCT---- 4075

Qy          2446 hrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnP 2466
                                        |||            :::||||||   |||   ||||
Db          4074 -----------------------CCACCACCAACCGCCACAACGCCT---CCGGCGAACC 4041

Qy          2466 roLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro------ProProGlyL 2484
                 ||                     |||            ||||||         |||||||
Db          4040 CA------CCCCACCGGGGAACCGGCACCCCATACCCACCACCGCGATCGGCTCCTGGGC 3987

Qy          2484 euProAlaGlySerGlyProLeuAlaGly 2493
                 ||   :::   |||   |||:::   |||
Db          3986 TTCGAAGCTCTTCTGCGCCGGTGCGAGGA 3958

RESULT 15
US-08-804-227C-13/c
; Sequence 13, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13987 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  350..13987
US-08-804-227C-13

Alignment Scores:
Pred. No.:              1.01e-20      Length:        13987
Score:                  652.00        Matches:       648
Percent Similarity:     30.21%        Conservative:  260
Best Local Similarity:  21.56%        Mismatches:    1034
Query Match:            4.93%         Indels:        1081
DB:                     2             Gaps:          133
```

```
US-09-522-753-5 (1-2517) x US-08-804-227C-13 (1-13987)

Qy        15 ThrGluProArgTyrProProHis---------------------------SerLeuSer 25
             |||:::      :::   ||||||                              ::::::
Db     10684 ACCCAGACCACGCACCACCCCCACCCCATCCGCGAAACGCACCGCCTCCCGCGCATGACG 10625

Qy        26 TyrProVal------------GlnIleAlaArgThrHisThrAspValGlyLeuLeuGlu 41
             :::|||:::            |||:::   |||||||||
Db     10624 CACCCAATACCCCGGATCCCCCAACTCCCCCGAACCCACCACCCCACCCGACACACCCGA 10565

Qy        42 TyrGlnHisHisSerArgAspTyrAlaSerHisLeu------------SerProGlySer 57
             :::   ||||||:::|||   :::   :::   |||            :::|||
Db     10564 CACCACCACCACACCCGGCCGCACCCGACCGAACTCCAACGACTCCACAACCCCCCGGAA 10505

Qy        58 IleIleGlnProGlnArgArg------------ArgProSerLeuLeuSerGluPheGln 73
             :::      ||||||   |||            |||         :::         |||
Db     10504 CTC------CCCCAACACCGGCTCCATCAACACCGAATGAAACCCATGCGACACATCCAA 10451

Qy        74 Pro------------------------------GlyAsnGluArgSerGlnGluLeuHis 83
             |||                                 |||   |||   |||      |||
Db     10450 CCGCCGGCACTCCACCCCATCCCCCATCAGCGTGACAACCACCGACTCCAGCACACCCAC 10391

Qy        84 LeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIle 103
             :::      :::   |||         |||
Db     10390 ATCACCCGACAACACCACCGACCGCGGCCC------------------------------ 10361

Qy       104 GluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgProSerProLeu 123
                   :::|||   |||      |||   |||   |||         |||:::|||
Db     10360 ATTCACCGCCGCAACCGACACCCACTCCC-CCAACCCCTCAACAACCCCCCGCACCACCG 10302

Qy       124 LeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThrGly 143
                   |||      |||      |||                           |||
Db     10301 ACTCCGACGCCCCCACCGACCACATCCCCCCACCCACCGGCAACCCACCCATCAACC--- 10245

Qy       144 LysLeuGluProValSerProProSer---------------------------ProPro 154
                      |||      ||||||:::                           ||||||
Db     10244 ------CACCCCGCGCCACCACCAACCGCACCGCATCACCCAACGACAACACCCCCGCCA 10191

Qy       155 HisThrAsp------------ProGluLeuGluLeuValProPro-ArgLeuSerLysGl 170
             ||||||               |||      :::      ||||||    |||
Db     10190 CATACGCAGCAGCCACCTCCCCCACCGAATGACCCAACACCACCGACACCTCCACACCCC 10131

Qy       170 uGluLeuIleGlnAsnMetAspArgValAspArgGluIleThrMetValGluGlnGlnIl 190
                 |||   |||      ::::::      |||   :::         :::::::::::
Db     10130 GAGCCTC---CAACGCCCGGAACAA------CGCCACCTCCAACGCAAACAACGCAGGCT 10080

Qy       190 eSerLysLeuLysLysLysGlnGlnGlnLeuGluGluGluAlaAlaLysPro-------- 207
             :|||||||||:::      ||||||      :::         :::   |||
Db     10079 GAGCAAACTCCGTCCGACCCAACAACCCACCCAAAACCGAGTCAGCAT-CCGCGAACATC 10021

Qy       208 ----------------------ProGluProGluLysProValSerPro----------- 216
                                   |||   |||      |||   :::|||
Db     10020 ACATCCCGCAACGACCGACCATCCACCTCCCCCACCACCGACAACACCTCATCCAGCACC 9961

Qy       217 ----------------ProPro-IleGlu---SerLysHisArgSerLeuValGlnIleI 230
                             |||||| |||:::   :::   |||   :::   ::::::   :
Db      9960 TCCGCAAACACCCCACCCCCCGCATACAACCCACGCCCCATACCAACCCACTGCGCACCC 9901

Qy       230 leTyrAspGluAsnArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyP 250
             ::      |||                        ||||||
Db      9900 TGACCCGTGAACAACACCCCCACCCCACGACCCCGCACCGC------------------- 9860

Qy       250 roGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHis------- 267
                           |||   ::::::               |||      |||
Db      9859 -----------ACTCCCCCGCACCACAGCACCGGCCCCCGAAGACACACCCACAGGAGCC 9811

Qy       268 --------GluAsnIleLysIleAsnGln------------------AlaMetArgLysL 279
                        :::|||:::      |||                  |||         :
Db      9810 CCGGCAGCCACCACATCCAGACCCGCCAGCAACTCCTCACGCCCCCGGCCCACCACCACC 9751

Qy       279 ysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheC 299
             ::   :::         :::   |||:::|||   |||
Db      9750 GCACGCTCACCGAACGACGCACGCGACACCACCAACGA---------------------- 9713

Qy       299 ysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnA 319
                                                :::   :::   ||||||      :
Db      9712 --------------------------------CCCCGCCACATCCCGCGGATCCGCACCC 9685

Qy       319 snProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProG 339
             ::|||               :::
Db      9684 ACCCGCTCCACACACTCACGCAACCGCAC------------------------------C 9655

Qy       339 luIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlyS 359
               :::|||   ||||||   |||:::   |||   :::
Db      9654 GCCTGCGCCCGCAACGCCCCCTCCGACCGCGCAGAAAGCAC------------------- 9614

Qy       359 erGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyL 379
                                    |||:::   |||   |||      :::
Db      9613 --------------CCACGACAACGGACCCGTCACCGCGTCGGGCCGCTC---------- 9578

Qy       379 euSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetL 399
                                       :::   |||   :::   |||:::   |||   |
Db      9577 --------------------------AGGTTCAGGGCCGTCCCCGTGCTCGCTCCCGGCC 9544

Qy       399 euTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspP 419
             ||
Db      9543 TC---------------------------------------------------------- 9542

Qy       419 roMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluT 439

Db      9542 ------------------------------------------------------------ 9542

Qy       439 hrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuG 459
                  |||   :::|||:::|||         :::   |||:::      :::      |
Db      9541 -----CGCCGGCGCTTCCTCCAG------GACCACATGGGCGTTGGTCCCGCTCACCCCG 9493

Qy       459 luArgLys-ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsn 478
             |||||::: |||   |||                        |||:::
Db      9492 AACGCCGACACCGCGGCCCGC------------------CGCACCCGCTCCACCCGCCGC 9451

Qy       479 TyrLysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGln 498
                         |||      |||         |||   :::|||            :::
Db      9450 GGCCACGACCGGGTCTCGGTCAGCAG-CTCCACCGCACCCGAAGCCCACTCCACCTTCGA 9392

Qy       499 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGlu 518
             :::::::::         |||:::   :::|||:::|||||||||
Db      9391 CGACGGCGCATCCACATGCAGCGTCCGCGGCAACGACCCATGCCGCATCCCCTGCAC--- 9335

Qy       519 GluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluVal 538
                            :::                                 |||:::|||
Db      9334 ---------CACCTTGATCACACCACCCACACCCGCAGCCGCCTGCGCATGCCCGATGTT 9284

Qy       539 GluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsn 558
                      :::
Db      9283 CGACTTCAACGA------------------------------------------------ 9272

Qy       559 AspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArg 578
                |||   :::   :::|||                           |||      |||
Db      9271 CCCGAGCCACAACGGATCGCCGACCCGCTCCCGCCCGTACGTGGCCAGCAGCGCACCCGC 9212

Qy       579 LysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrPro 598
                   |||||||||:::
Db      9211 CTCGATCGGATCACCCAGCGCCGTCCCCGTACCGTGCGCCTCCACCAC------------ 9164

Qy       599 GlnGlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGlu 618
                            :::         :::         |||   |||   |||:::
Db      9163 ---------ATCCACGTCGGCGGGCACCAGCCCCGCATCAGCCAGCGCCTCACGGATCAC 9113

Qy       619 GluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAla 638
                :::                  |||            |||            ||||||
Db      9112 CCGCTG------------------CTGCGCCGGACCACTCGGCGCCGTCAGCCCATTGCT 9071

Qy       639 ArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLys 658
             |||   |||
Db      9070 CGCACCGTC--------------------------------------------------- 9062
```

```
Qy      659 ArgGlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsn 678
                     :::|||               |||:::         :::
Db     9061 ------CTGATTGACCGCACTGCCCCGCACCACCGCCA-GCACCCGATGACCGTTGCGGC 9009

Qy      679 AlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProPro 698
            |||||||||      |||   ||||||   ||||||                  ||||||
Db     9008 GCGCGTCGG------AAAGCCGCTCCACCAGCAGCACACCCACAC---------CCTCCG 8964

Qy      699 ValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetVal 718

Db     8963 CCCACGCCGTCCCGTCCG------------------------------------------ 8946

Qy      719 GluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGly 738
                              |||   |||   |||   :::   |||         :::
Db     8945 ------------------CACCC-TCCGCGAACGCCTTGCACCGCCCGTCCGCCGCCAGC 8905

Qy      739 ProAlaThrValAsnAsnSerSerAspThrGluSerIle-ProSerPro----------- 754
            |||            ||||||::::::|||      ::: |||||||||
Db     8904 CCCCGCTGCCGCGAGAACTCCACGAACACCGTCGGCGTCGCCATCACCGTCACCCCACCG 8845

Qy      755 -------------HisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProPr 770
                         ||||||      |||      |||                  |||||
Db     8844 GCCAACGCCAGATCACACTCACCCCGCCGCAACGACTGCACCGCCAGATGCAACGCCACC 8785

Qy      770 oAlaThrLeuGlyAlaAspGlyProPro------ProGlyProProThrPro-------- 785
            |   |||               ||||||      |||   ||||||||||||
Db     8784 AACGACGACGAACACGCCGTGTCCACCGTCAGCGCCGGCCCCTCCAACCCCAACGCATAG 8725

Qy      785 ------------------------------------------------------------ 785

Db     8724 GCAACCCGACCGGAGATCACACTGCCGAGGGTCCCCGTGAGCACATGCCCGCCGTAGGTG 8665

Qy      785 ------------------------------------------------------------ 785

Db     8664 TCGGCGGCGTCCCCCAGCTGGGTCGCGTAGTCCTGCTGGGAGATGCCGCTGAAGACACCG 8605

Qy      786 ----------------------ProArgArgThrSerArgAlaProIleGluProThr-- 797
                                  |||   |||         ||||||      ||||||
Db     8604 GTCCGGGTGCCGCGCAGCGAGTCCGGAACGATGCCGGCCCGCTCCAGCGCCTCCCACGAC 8545

Qy      798 ----ProAlaSerGluAlaThrGlyAlaProThrPro----------------------- 808
                ||||||:::   |||         |||   |||
Db     8544 GTCTCCAGCAGCAACCGCTGCTGCGGGTCCATGGCCGTGGCCTCGCGCGGCGATATCCCG 8485

Qy      808 ------------------------------------------------------------ 808

Db     8484 AAGAACTCCGCGTCGAACTCGGCCGCGGAACGCAGGAACCCGCCTTCCCGTACATAGCTC 8425

Qy      809 -ProProAlaProProSerProSerAlaProProProValValProLysGluGluLysGl 828
                ||||||||||||:::|||            |||      |||         :::
Db     8424 TTGCCCGGCACCCCCGGCTCCGGGTCGTACAGCCCCTCGACGTCCCAGCCACGGTCGGCC 8365

Qy      828 uGluGluThrAlaAlaAla------------ProProValGluGluGlyGluGluGlnLy 844
                         |||               ||||||         |||
Db     8364 GGGAACGGGGAGAGCGTGTGCCCGTCACCGGCCACCAGGTCCCACAGGTCGTCCGGCGAC 8305

Qy      844 sPro-------------------------------------------------------- 845
             |||
Db     8304 GCCACACCGCCCGGGTAACGGCACGCCATGCCCACGATCGCGATCGGCTCACGCTCGTCG 8245

Qy      846 -------ProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLy 863
                   ||||||            |||      :::               |||
Db     8244 GCCATGACCGGCGTG-GGGGCGGCCGCTGCCGGCTCCGCCTCGCTGTCGGGACCGAACAG 8186

Qy      863 sSerGluCysThrGluGluAlaGluGlu-GlyProAlaLysGlyLysAspAlaGluAlaA 883
             |||   |||      ||||||   ||| |||   :::   |||:::   |||
Db     8185 TTCGTCGTGCAGGTGGAGGCGAGCGAGAGGGGGGGTCGGGTGGTCGAAGAGCA------- 8133

Qy      883 laGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgA 903
                 |||   :::   |||:::   :::   ::::::      |||   :::   |||
Db     8132 -----GCGTGTTCGGCAGGTTCAGGCCGGTGGCAGCGGTCAGGCGGTTACGCAGTTCGAG 8078

Qy      903 laThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerA 923
              |||      :::|||:::   |||   :::   :::   |||   :::|||   ||||
Db     8077 A-ACGGTGAGGGAGTCGAACCCGGCGGACCGGAAGGCCAGCTCCGGGTCGACCGCGTCGG 8019

Qy      923 laAsp------------------------------------GluValAspGluAla---- 929
            ||:::                                       |||      |||
Db     8018 CCGAGGCGTGGCGCAGGATCCCGGCGGCCTCCGAGCGCACAAGCGTGAGCAGAGCACGAC 7959

Qy      930 --------------GluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuL 945
                          ||||||      ::::::                  ||||||
Db     7958 GGCGTTCCGCCGGCGAGGGGCCGACGGAGAGCCACGGCCGGCCCTCGCCCCCGTCCTGTC 7899

Qy      945 euThrProThrGlyAspPro---------------------------------------- 951
                 |||   |||   |||
Db     7898 CGCTCCCGCCCGGTGCGCCGTCGCCCGGCCGCGTCCCTGGGGTGAGAGCGGCACGGTCGT 7839

Qy      952 --------------------------ArgAlaAsnAlaSerProGlnLysProLeuAspL 963
                                      ||||||      :::|||   |||         :
Db     7838 CCGCGGCATGGCCGGCACGCGCCGAACGGGCGGCGGAGACGCCGTCGAAGAGCGACCAGG 7779

Qy      963 euLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisG 983
            ::::::::::::  :::|||               |||   :::|||::::::
Db     7778 TACGGCGGGTGACCGAACGC---------------CCGACGAAGGTCTCCCAGTCCACGT 7734

Qy      983 luProProArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProG 1003
                          :::   :::|||            |||   |||      ||||||
Db     7733 CGGCGACGAGCAGGGAGCTCTCGCCGGCCGCGTGCGTCCCTGCGGCAGAGGTCCCGCCG- 7675

Qy     1003 lnAsnLeuGlnProGluSerAspAlaProGlnGlnProGly------SerSerProArgG 1021
                                   ||||||                  ||||||   |||
Db     7674 -----------------------GCGCCCACCACGTCGTCGAGGACGTCGAGCGCGCGCT 7638

Qy     1021 lyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAla------GluA 1039
                 ||||||:::      |||         |||:::   |||||||||      ||||
Db     7637 CGGGGTCCAGGGCCGTGACGCCGTCCCGTTCGAAGGAC---TTCGCCGCCGTTCCGGAGG 7581

Qy     1039 laGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProP 1059
            ||         |||         |||                        |||:::
Db     7580 CCATTCCGGCACCGGCCCACGGCCCCCAGGCGATCGAGGCGGCACGGGCTCCCCTGGCCC 7521

Qy     1059 roArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaP 1079
              |||         :::||||||      |||
Db     7520 GGCGCCGCTGCGCCAGGGCGTCGAGCGCGGCG---------------------------- 7489

Qy     1079 roProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgP 1099
                                |||            ||||||   |||   ||||||
Db     7488 --------------------TTGGCGGCCGCGTAGACGGCCTGCCCTCCGTTGCCCCACA 7449

Qy     1099 roProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuG 1119
              ||||||:::                     ||||||   :::
Db     7448 CGCCGACGATGGAGGAGAAGAGCACGAAGGCGTCGAGCGGAAGGTCCGCGGTCAGCTCGT 7389

Qy     1119 luArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHis------------- 1134
              |||         |||::::::         |||      :::|||
Db     7388 CCAGGTGACGGGCGGCGGTGACCTTGTCCCGCAGTACGGC-GTCCACCCGGTCGGAGGTG 7330

Qy     1135 --------------------------------------ValProTyrSerGluHisAlaL 1142
                                                  |||   |||   :::
Db     7329 AGCGAGTCGAGCACACCGTCGTCGAGCACACCTGCCGCGTGCAGTACGGCGACGAGGGGC 7270

Qy     1142 ysAlaProValGly------------------ProValThrMetGly------------- 1151
              ||||||   |||                  ||||||      |||
Db     7269 C-GCTCCTCGGGGAGGCCCGCGAGGAGCGCGGCCAGTGCGTCCCGGTCGGCGACGTCGCA 7211

Qy     1152 --------LeuProLeu------------ProMetAspProLysLysLeuAlaProPheS 1165
                       ||||||            |||      |||:::::::::   |||
Db     7210 GGCCTCGACGGCCACTTCGGAACCGAGCGCCGTCAGCTCCTCGACGAGTTCGCCCACTCC 7151

Qy     1165 erGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProPro-GluSerLeu 1184
              |||                     ||||||   :::|||||||||||| |||
Db     7150 GGGGC------------------CGGCCGCGC---CGCGCCGGCCCACCAGGAGCAG--- 7115

Qy     1185 GlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeuGlySerValPro 1204
                                             ||||||         |||
Db     7114 ------------------------------ACGCGCGGCGCCCGACCGGGCGAGCCTGCG 7085

Qy     1205 GlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp-SerAlaIleTh 1224
            |||         |||:::|||         |||      |||   ||| |||   :::
Db     7084 GGCGACCTGCGCACCCAGGGC---------ACCGGTGCCCCCGGTGATCAGTACGGTG-- 7036

Qy     1224 rTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleTh 1244
                                        |||   |||            ||||||   ||
Db     7035 ----------------------------CCGGAGGACCACCGGTCGTCCGGCACGACCAC 7004

Qy     1244 rArgIleIleGly------GluAspSerProSerArgLeuAspArgGlyArgGluAspSe 1262
            |      :::               ||||||:::   :::      ||||||   |||
Db     7003 GTCCGAGGTGTCGTTCCGTGCGTCCTCGCCGACGGCGGTGGC-GGGGGCCGTGGTGACGG 6945

Qy     1262 rLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGl 1282
                      |||            :::|||:::      |||            :::
Db     6944 TGCGAGCGGAGGCCGG---GACGGGGACGGGCGCGGGAACCACACGGCGGACATGGACGC 6888

Qy     1282 yGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHi 1302
             |||   :::|||                        |||:::   |||   |||
Db     6887 CGGAGG-GCCGTACCG---------------------CGAACTGGTTCTCTCCCC----- 6855

Qy     1302 sGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIl 1322
                   |||||||||:::|||||||||                        ||||||
Db     6854 -CGGCGAGCAGCGCCCCGACGAACGTACGGA---------------GCACCGCGGGCGAC 6811

Qy     1322 eSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerPr 1342
                |||                        ||||||   ||||||   |||   |||
Db     6810 GCACTCC------------------------CGGGCAGGTCCGCCAAGCCGCCCCAGCGC 6775

Qy     1342 oHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSe 1362
                   ::::::   :::      :::|||                  :::|||
Db     6774 CGCCCCAGTTCGAGCGCGGCGACCTGTCCGA------------------GTCCCC----- 6738

Qy     1362 rTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArgGluGl 1382
                                                 |||:::      :::|||:::||
Db     6737 ----------------------------------ACAGCTGAGCGCCCCCGACGGAAGGG 6712

Qy     1382 yThr-------------------------------------------------------- 1383
            |:::
Db     6711 GTCTCGTCCGGCCGCACGGCCGCCGCCTCACGGGTGGTGATCCAGAGCGGGGCTTCGGGT 6652

Qy     1384 -ProProProProProProSerArgAsp-------------------------------- 1392
             ||||||   |||   |||   |||
Db     6651 ACCGCCTCGCCCAGCGCCTGGACGAGGGCGACGGCGTCGGCCGCGTCCGGCTCGGCCGGG 6592

Qy     1393 ----LeuThrGlu---------------------AlaTyrLysThrGlnAlaLeuGlyPr 1404
                   ||||||                        ::::::      |||   |||||
Db     6591 GCGAGCCACAGAACTCGGGTGCTGCGGGCGGCACCGTCACCGGCGGCCCGCAACAGGTCC 6532

Qy     1404 oLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu----------- 1420
            |   :::|||   |||         ||||||            :::
Db     6531 C---CGGCTGCCTCCC---------GGCCTGTACCGACCGGCACGGCGACGGTGTGGACC 6484

Qy     1421 -------AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGl 1438
                         |||   :::|||         |||      :::|||      |||
Db     6483 TCGGCCCCGCGCACGGCGAGTTCACGTGCGATCGCGGGGGCGAGTCCGGAGTCGGCCG-- 6426

Qy     1438 uLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLy 1458
                         |||||||||                        |||
Db     6425 ----------TGTCCTCGTCCG------------------------GGGCCAGAAGCAGC 6400

Qy     1458 sTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGl 1478
             :::|||      |||               ::::::                     ||
Db     6399 ACATGACCGGCAGGCCCGGGGTTGCCGCGGTCGACGACGGCCGGCTGCCAAGCGACGCGG 6340

Qy     1478 ySer---------ProGlyArgThrPheProProValHisProLeuAspValMetAlaAs 1495
            ||||         |||      |||   |||         |||            |||
Db     6339 TAGCGGATCCCGGCCGTCGTCGACCGTCCCCGCTCAGTGGCCCGAGCGTCCGCCGGCGAG 6280

Qy     1495 pAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAl 1515
                ||||||      |||               |||   :::         |||   ||
Db     6279 CCGGCGGGCTCGACGCCGG---------------TCACGGCGGTGCGAACGGGGTCCAGC 6235

Qy     1515 aSerSerSerGlyGlySerIleAlaArgGly----------------------------- 1525
            |||||||      |||   :::   ||||||
Db     6234 CAGTAGC------GGTCGCGTTGGAAGGGGTAGGTGGGGAGGTCGATGCGCCGGCCGGTG 6181

Qy     1526 ----------------------------------AlaProValIleValProGluLeuGl 1534
                                                 |||         |||
Db     6180 CTCCCGGTGTGGAGTGCCGTGGCGTCCAGGCCGGCGTCCCGGGTGAACACCGTCGCCAGC 6121

Qy     1534 yLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLe 1554
                ||||||      |||                              |||   |||
Db     6120 GCCGCCTCGAACACCTCCC------------------------------GCTCCGCACGG 6091

Qy     1554 uProArgGlySerProValThrMetArgGluProThrProArg-LeuGlnGluGlySerL 1574
             |||      ||||||               |||   |||    :::
Db     6090 CCCCGCCGCATCGCCGGCA------------CCACCACCACATCATC------------- 6056

Qy     1574 euSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGlu------- 1591
                             :::      :::   ||||||      ||||||:::
Db     6055 -----------------ACCGGCCCCCAGGCACTCACCCGCCATCCCCGTCAGCACCCCA 6013

Qy     1592 -----------IleAlaLysSerProHisSerThrValProGluHisHisProHisProI 1608
                       :::   :::   ||||||   |||         ||||||||||||||||
Db     6012 TGCGGACCCACCTCCACCAACGTCCCCAC---ACCCAGACCACGCACCACCCCCACCCCA 5956

Qy     1608 leSerProTyrGluHisLeuLeuArgGlyValSerGlyVal------------AspLeuT 1624
            ||               |||   |||            :::               |||
Db     5955 TCCGCGAAACGCACCGCCTCCCGCGCATGACGCACCCAATACCCCGGATCCCCCAACTCC 5896

Qy     1624 yrArgSerHisIleProLeuAlaPheAspProThrSerIleProArg------------- 1639
              |||:::|||   |||            ||||||      |||
Db     5895 CCCGAACCCACCACCCCACCCGACA-CACCCGACACCACCACCACACCCGGCCGCACCCG 5837

Qy     1640 -----------------------------GlyIleProLeuAspAlaAlaAlaAlaTyrT 1650
                                         |||   |||         |||   :::
Db     5836 ACCGAACTCCAACGACTCCACAACCCCCCGGAACTCCCCCAACACCGGCTCCATCA---- 5781

Qy     1650 yrLeuProArgHisLeuAlaProAsnProThrTyrPro-HisLeuTyrProProTyrLeu 1669
                                ||||||   |||:::    |||:::   ||||||   |||
Db     5780 -----------------ACACCGAATGAAACCCATGCGACACATCCAACCGCCGGCACTC 5738

Qy     1670 ---------------IleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThr 1684
                              ||||||   |||   |||   |||
Db     5737 CACCCCATCCCCCATCAGCGAGGCAA-CCACCGACTCCAGCA------------------ 5697

Qy     1685 IleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAla 1704
                                             ||||||   |||            |||
Db     5696 ---------------------------CACCCACATCACCCGACAACACCACCGACCGCG 5664

Qy     1705 GlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyr 1724
                  :::
Db     5663 GCCCATTCA--------------------------------------------------- 5655

Qy     1725 AlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuVal 1744
                                                   |||      |||
Db     5654 ---------------------------------------CCGCCGCAACCGACACCCACT 5634

Qy     1745 ProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAla 1764
            ||||||||||||      |||                                 |||:::
Db     5633 CCCCCAACCCCTCAACAACCCCCCGCACCACCGACTCCGACGCCCCCACCGACCACATCC 5574

Qy     1765 ProGlnProPheSerSerArgHisSerSerSerProLeuSerProGlyGlyProThrHis 1784
            |||   |||   ::::::      |||:::   |||      |||      ||||||
Db     5573 CCCCACCCACCGGCAACCCACCCATCAACCCACCCCGCGCCACCA------CCAACCGCA 5520

Qy     1785 LeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArg 1804
                     ||||||||||||               ::::::
Db     5519 CCGCATCACCCAACGACAACACCCCCGCCACATACGCAGCAGCCACCTCCCCCACCGAAT 5460

Qy     1805 AspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAla 1824
            |||                                 :::      |||      |||
Db     5459 GAC------------------------------CCAACACCACCGACACCTCCACACCCC 5430
```

```
Qy        1825 ProIleTrpArgProGlyThr---------------------------GluGlnSerSer 1835
                           |||||||||                              :::   :::
Db        5429 GAGCCTCCAACGCCCGGAACAACGCCACCTCCAACGCAAACAGCGCAGGCTGAGCAAACT 5370

Qy        1836 GlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSer 1855
                  |||      ::::::               :::   :::||||||      |||
Db        5369 CCGTCCGACCCAACAACCCACCCACAGAACCGGAACCAGAACCGACCC---CAGCACCCG 5313

Qy        1856 HisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgPro 1875
               |||   |||   |||   |||      ||||||               ::::::|||
Db        5312 CACCGGCACCCGCATCAGCCCCGGCACCCGCGTCCA------------CGTCGACGTCGC 5265

Qy        1876 SerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysPro 1895
                        |||                        |||         |||      |||
Db        5264 CGAACATCACATCCCGCA------------------ACGACCGACCATCCACC-TCCCCC 5224

Qy        1896 ThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProPro 1915
               |||:::      ::::::|||::::::|||         |||         :::   |||
Db        5223 ACCATCGACAACACCTCATCCAGCACCTCCGCAAACACCCCACCCCCGCATACAACCCA 5164

Qy        1916 ---------AlaThrHisCys---------ProLeuGlyGlyThrLeuAspGlyValTyr 1929
                           |||||||||         |||:::      |||
Db        5163 CGCCCCATACCAACCCACTGCGCACCCTGACCCGTGAACAACACC--------------- 5119

Qy        1930 ProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGlu 1949
               ||||||                                 |||      |||||||||:::
Db        5118 CCCACC------------------------------CCACCCGGCACAGCCCGCCCGGAC 5089

Qy        1950 ArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeu 1969
               |||         ||||||
Db        5088 CGCACACCCCCAGACACC------------------------------------------ 5071

Qy        1970 GluProAlaSerSerProSerLysGly------SerGluProArg---Pro-LeuValPr 1986
                  ||||||::::::|||   :::         |||   ||||||   |||       ||
Db        5070 CTCCCGGCAGCAAACCCACCAAGAGCAGCGACGAGCTCCCCACGCCCACCACACGGCACC 5011

Qy        1986 oProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHi 2006
               ||||                     |||      ||||||            ||||||
Db        5010 ACCGCACGGTGCTCGAAAAGAGTCCGCGTCGAAGCCAGCGAATAAG------CCACATCC 4957

Qy        2006 sAlaSerProAspProPro---------------------AlaProProAlaSerAlaSe 2019
                |||   |||   ||||||                     ||||||||||||   |||::
Db        4956 CGCGGATCCGCACCCACCCGCTCCACGTACTCACCCAACCGCACCGCCTGCGCCCGCAAC 4897

Qy        2019 rAspProHisArgGluLysThr-GlnSerLysProPheSerIleGlnGluLeuGluLeuA 2039
               :   |||         :::||| |||   :::                           |
Db        4896 GCCCCCTCCGACCGCGCGGACAACAACAGCGG------------------------CGAC 4861

Qy        2039 rgSerLeuGlyTyrHisGlySerSerTyr---SerProGluGlyValGluProValSerP 2058
               |||||            |||::::::      |||   :::   |||:::   :::   |
Db        4860 GGAGCAGCCGCGTCCGCGGTACCGCGGTCAGGAGCAGCGACGCCGTCGATGTCCTCGACC 4801

Qy        2058 roVal------------------------------------SerSerProSerLeuThrH 2066
               ||:::                                    :::   |||         |
Db        4800 CACTCCGGCGCCTCCTCGATGATCAGATGCGCGTTCGTCCCACTGATCCCGAACGACGAC 4741

Qy        2066 isAspLys------------------------GlyLeuProLysHisLeuGluGluLeuA 2078
               ||   :::                        |||   |||      |||   :::
Db        4740 ACAGCCGCCCGGCGCGGGCGGTCCGACCGCCGGGGCCACCGGCGCGCCTCTGCCAGCAGG 4681

Qy        2078 spLysSerHisLeuGluGlyGluLeuArgPro------LysGlnProGlyProValLys- 2095
                             :::|||      ||||||         :::   |||||||||:::
Db        4680 GC------------GACGGTGCCCGACGACCAGTCGACGTGCGTGCTGGGCCGGTCCGCG 4633

Qy        2096 --------------------------LeuGlyGlyGluAlaAlaHisLeuProHisLeuA 2107
                                         |||   |||:::||||||      ||||||   |
Db        4632 TGAAGTGTCCGCGGCAGCGTGCCGTGCTCCA-GGGCAAGCAGCATCTTGACCACACCCGC 4574

Qy        2107 rgProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValL 2127
               |||||                                             |||
Db        4573 GGCCC---------------------------------------------CCGCGGCGGC 4559

Qy        2127 ysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspT 2147
```

```
                 |||      |||      |||   |||
Db        4558 CTGGGTGTGACCGACGTTGGACTTCAGCGACCCGAGCC---------------------- 4521

Qy        2147 yrThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheP 2167
                 ||||||      |||                     ||||||         ::::::|
Db        4520 --ACAAGGGGTCGCCCA---------------------CCCGCTCACGCCCGTACGTGGC 4484

Qy        2167 roGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProP 2187
               ||   |||      |||      |||   |||      ||||||            ||||
Db        4483 CAGCAGCGCGCTCGCCTCGATCGGAT---CGCCCAGCGCCGTCCCCGTACCGTGCGCCTC 4427

Qy        2187 roProAspHisGlyAlaProAlaArgGly------------------------------S 2197
               |||||            ||||||||||||
Db        4426 CACCGCGTCCACGTCACCCGGCGCGAGGCCGGCGTCGGCCAGGGCAGCTCGGATGACGCC 4367

Qy        2197 erProHisSerGluGlyGlyLysArgSer------------------------------- 2206
                 |||         |||   :::||||||
Db        4366 CTCCTGCGCGGAGCGGTTGGGAGCGGTCAGGCCGTTGCTGGCACCGTCCTGGTTGACCGC 4307

Qy        2207 -----ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProV 2225
                    |||   |||
Db        4306 ACTGCCCCGCACCACGC------------------------------------------- 4290

Qy        2225 alSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuL 2245
                    |||      |||:::         |||      |||:::|||      |||:::
Db        4289 -----CCAGCACCCGGTGTCCCGCACGCCGGGCGTCGGAGAGCCGCTCCAGTGCCAGTAC 4235

Qy        2245 euTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnT 2265
                             :::|||      |||            |||   |||      |||
Db        4234 G-GCGACACCCTCGGACCAGGCCGTACCG------------TCGGCATCGGCGGAGAACG 4188

Qy        2265 hrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysS 2285
                    |||   |||:::      |||               ||||||
Db        4187 GCATGCAGCGGCCGTCGGGGGCGAGCCCGCGCTGCCGCGAGAACTCCACG---------- 4138

Qy        2285 erLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrA 2305
                                                     |||         |||:::
Db        4137 --------------------------------------AACATGCCGGGCCCCGACATCA 4116

Qy        2305 snIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuM 2325
                 ::::::   |||               :::   |||
Db        4115 CCGTGGCGCCGCCCGCCAGAGCCAGCCCGCACTCGCCCCGG---------CGCAGCGCCT 4065

Qy        2325 etThrTyrArgSerGlnAlaVal------GlnGluHisAlaSerThrAsnMetGlyLeu- 2342
                 |||   |||      |||         :::|||||||||   :::   :::
Db        4064 GCACGGCCAGGTGCAGCGCCACCAGGGACGAGGAGCACGCGGTGTCGACGGTGAGCGCCG 4005

Qy        2343 -----------------------------------GluAlaIleIleArgLysAlaLeuM 2351
                                                  ||||||   :::            :
Db        4004 GTCCGCCGGTGCCGAGGGTGTAGGCGATCCGGCCGGAGGCCACGCTCGCGGTGGTGCCGG 3945

Qy        2351 etGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProL 2371
               ::   :::|||      :::      :::|||      |||         ::::::
Db        3944 TGAGCAGGTACCCCTCGTATCCGTCCGCTCCCTCGTGAAGCCTCGGCCCGTACTCCGGTG 3885

Qy        2371 euAsnAlaSerAlaSerLeuProAlaAlaMetPro---------------IleThrAlaA 2386
                    |||      :::   ||||||:::   |||                  |||   |
Db        3884 CGGTGGCGCCGACGAACACTCCGGCGTCGCCGCCGCGCAGTGACTCGGGGAGCACTCCGG 3825

Qy        2386 laAspGlyArgSerAspHis-ThrLeu------------------ThrSerProGlyGly 2399
               ||      |||:::   ||| ||||||                        |||
Db        3824 CGTCCTCGAGGGCCTCCCACGACGCTTCCAGCAGCAAGCGCTGCTGCGGGTCCATGGCCG 3765

Qy        2400 GlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro 2419
                     :::|||   :::|||   |||   :::   |||:::   :::   |||
Db        3764 TGGCCTCGCGCGGCGATATCC---CGAAGAACTCCGCGTCGAACTCGGCCGCCG------ 3714

Qy        2420 GlyLeuAlaSerGlyAspArgProPro---------SerValSerSerValHisSerGlu 2436
                     ::::::|||   |||||||||         |||      :::      :::
Db        3713 ------AGTGCAGGAACCCGCCCTCCCGCACATAGGTCTTGCCGGGCGTCCCGGGCTCGG 3660

Qy        2437 GlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSer 2456
               |||         ||||||   |||            |||   :::|||      :::|||
Db      3659 GGTCGTACAGCCCGTCGAGGTCCCAG-CCGCGGTCGGTGGGGAAGGACGAGATGGCGTCG 3601

Qy      2457 AlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMet 2476
             |||                        |||               |||
Db      3600 GCG------------------------CCGGACTCCAGCAGCTCCCAGAGAGCGGCGCGC 3565

Qy      2477 AlaSerProProPro------------------ProGlyLeuProAlaGlySerGlyPro 2490
                |||   ||||||                  |||   :::      ||||||
Db      3564 GAGTCGACGCCGCCGGGGAGCCGGCAGCCGACACCCACGATGGCGATCGGGTCGTCGTCG 3505

Qy      2491 LeuAlaGlyProHisHisAlaTrpAsp---------------GluGluProLysPro--- 2504
                      |||      |||   |||               :::|||      |||
Db      3504 TCCCGGCTGCCGCCCCGGGCGTCCGACACCCCTACCGCGGCCGACGAGAGCACTCCGGGC 3445

Qy      2505 ---------------LeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
                            |||   |||:::         |||      :::   |||
Db      3444 CCGCCGGCCTCGAAGCTCTTCTGCGCCGGTGCGAGGACGGAGCCCGAGGGGGAG 3391

Search completed: November 12, 2005, 14:54:06
Job time : 41449 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         November 11, 2005, 09:34:07 ; Search time 35711 Seconds
                                     (without alignments)
                                     3415.244 Million cell updates/sec

Title:          US-09-522-753-5
Perfect score:  13215
Sequence:       1 MSGSTQLVAQTWRATEPRYP..........WDEEPKPLLCSQYETLSDSE 2517

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters:      9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09522753/runat_02112005_173640_20793/app_query.fasta_1.2695
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09522753_@CGN_1_1_14601_@runat_02112005_173640_20793 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                    Description
 ---------------------------------------------------------------------------
     1   13215   100.0    8561  9  AF113003             AF113003 Homo sapi
     2   12978    98.2    8686  9  AF125672             AF125672 Homo sapi
     3   12922    97.8    9053  6  AR447713             AR447713 Sequence
     4   11067    83.7    8388  10  AF113001             AF113001 Mus muscu
     5 10832.5    82.0    8544  10  AF125671             AF125671 Mus muscu
     6    9894    74.9    7465  10  AF113002             AF113002 Mus muscu
     7  9776.5    74.0    6339  6  CQ722208             CQ722208 Sequence
     8    7885    59.7    5989  9  HSU37146             U37146 Human silen
     9    4612    34.9    2930  9  S83390               S83390 T3 receptor
    10    4194    31.7    7949  9  AB028970             AB028970 Homo sapi
    11    4187    31.7    7780  10  MMU35312             U35312 Mus musculu
    12  4147.5    31.4    7940  6  CQ847916             CQ847916 Sequence
    13  4147.5    31.4    7940  6  AX578069             AX578069 Sequence
    14  4147.5    31.4    7940  9  AF044209             AF044209 Homo sapi
    15    4043    30.6    8018  9  AF087856             AF087856 Homo sapi
    16  4009.5    30.3    8959  5  AF495886             AF495886 Xenopus l
c   17  3736.5    28.3  191100  10  AC139377             AC139377 Mus muscu
c   18  3479.5    26.3  254449  2  AC097560             AC097560 Rattus no
c   19    3445    26.1  218129  2  AC136560             AC136560 Rattus no
    20  3344.5    25.3  219339  2  AC121005             AC121005 Rattus no
    21    3178    24.0    2842  9  BC004326             BC004326 Homo sapi
    22    3012    22.8    2964  10  BC047524             BC047524 Mus muscu
    23  2910.5    22.0    6300  6  CQ724013             CQ724013 Sequence
c   24  2891.5    21.9  205283  9  AC073916             AC073916 Homo sapi
    25    2187    16.5    3120  9  AF303586             AF303586 Homo sapi
    26    2155    16.3    1917  5  BC054296             BC054296 Xenopus l
    27    2147    16.2    1896  5  BC082706             BC082706 Xenopus l
    28  2091.5    15.8    6328  10  AB093281             AB093281 Mus muscu
    29    1900    14.4    6541  9  AB019524             AB019524 Homo sapi
    30    1839    13.9    4285  10  MUSRIP13G            L78294 Mus musculu
    31    1687    12.8    1891  9  BC050594             BC050594 Homo sapi
    32  1663.5    12.6    1850  9  BC058511             BC058511 Homo sapi
    33  1645.5    12.5    1808  9  BC068996             BC068996 Homo sapi
    34    1633    12.4    1741  9  BC056862             BC056862 Homo sapi
    35  1628.5    12.3    3025  5  BC049302             BC049302 Danio rer
    36    1611    12.2    2145  5  BC073704             BC073704 Xenopus l
    37  1530.5    11.6    1857  5  BC073253             BC073253 Xenopus l
    38    1396    10.6  161970  2  AC027706             AC027706 Homo sapi
    39    1391    10.5    2914  6  CQ850622             CQ850622 Sequence
    40    1391    10.5    2914  9  AK127788             AK127788 Homo sapi
    41  1350.5    10.2    1152  5  AY498876             AY498876 Xenopus l
    42  1172.5     8.9    3997  9  AF303585             AF303585 Homo sapi
c   43    1162     8.8  133947  5  AL590153             AL590153 Zebrafish
    44    1113     8.4     650  6  AX677743             AX677743 Sequence
    45    1099     8.3   11296  3  AF175223             AF175223 Drosophil


                              ALIGNMENTS


RESULT 1
AF113003
LOCUS       AF113003                8561 bp    mRNA    linear   PRI 20-MAR-1999
DEFINITION  Homo sapiens silencing mediator of retinoic acid and thyroid
            hormone receptor alpha mRNA, complete cds.
ACCESSION   AF113003
VERSION     AF113003.1  GI:4454551
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8561)
  AUTHORS   Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and
            Evans,R.M.
  TITLE     Unique forms of human and mouse nuclear receptor corepressor SMRT
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)
  MEDLINE   99178941
   PUBMED   10077563
REFERENCE   2  (bases 1 to 8561)
  AUTHORS   Downes,M.R., Ordentlich,P. and Evans,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk
            Institute for Biological Studies, 10010 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1..8561
                     /organism="Homo sapiens"
```

AAD20946

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /map="12q24"
                     /tissue_type="pituitary"
     CDS             2. .7555
                     /function="transcriptional co-repressor"
                     /note="hSMRT alpha; longer isoform than previously
                     reported"
                     /codon_start=1
                     /product="silencing mediator of retinoic acid and thyroid
                     hormone receptor alpha"
                     /protein_id="AAD20946.1"
                     /db_xref="GI:4454552"
                     /translation="MSGSTQLVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQH
                     HSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEF
                     IESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPEL
                     ELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPP
                     IESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAM
                     RKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRAKESKVREYYE
                     KQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQL
                     AVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKN
                     FGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQ
                     QQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVAS
                     KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETAK
                     KGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRK
                     KKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPA
                     TVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTP
                     ASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTG
                     KAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQD
                     SDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAA
                     IPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPA
                     DKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGL
                     HDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVG
                     PVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVP
                     GGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDS
                     LPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGR
                     AISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLL
                     KREGTPPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHT
                     PELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMA
                     DARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAG
                     HLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISP
                     YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPY
                     LIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAA
                     GPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTH
                     LTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG
                     GGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRS
                     TSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHA
                     FLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPA
                     SASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHL
                     EELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRV
                     VTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDH
                     GAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRD
                     GEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNI
                     SQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLS
                     ANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLA
                     SGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPP
                     PPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE"
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-258       Length:         8561
Score:                  13215.00         Matches:        2517
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-09-522-753-5 (1-2517) x AF113003 (1-8561)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGTCGGGCTCCACACAGCTTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 61
```

```
Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 121

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCGGGCTCCATCATCCAG 181

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 241

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 301

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 361

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 421

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 481

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 541

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 601

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 661

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 721

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        722 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 781

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        782 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 841

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        842 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAAGCAGAAGTTCTGCCAG 901

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        902 CGCTATGACCAGCTCATGGAGGCCTTGGAAAAAAAGGTGGAGCGCATCGAAAACAACCCG 961

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        962 CGCCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAAAAGCAGTTCCCTGAGATC 1021

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1022 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1081

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1082 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1141
```

```
Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1142 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1201

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1202 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1262 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1321

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1322 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1381

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1382 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1441

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1442 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1501

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1561

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1562 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1621

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1622 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1681

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1682 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1741

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1742 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1801

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1802 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1861

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1862 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1921

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1922 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1981

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1982 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2041

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2042 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2101

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2102 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2161

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2162 GCTGAAGCCTTACATGCCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCCCAGCC 2221

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2222 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2281

Qy          761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2282 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2341

Qy          781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2342 GGCCCACCCACCCCACCACGGAGGACATCCCGGGCCCCCATTGAGCCCACCCCGGCCTCT 2401

Qy          801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2402 GAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCTCCTCCT 2461

Qy          821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2462 GTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2521

Qy          841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2522 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2581

Qy          861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2582 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2641

Qy          881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2642 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2701

Qy          901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2702 GGCAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2761

Qy          921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2762 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2821

Qy          941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2822 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2881

Qy          961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2941

Qy          981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2942 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3001

Qy         1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3002 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3061

Qy         1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3062 GGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCCTTCGCAGCCGAGGCCCAG 3121

Qy         1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3122 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 3181

Qy         1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3182 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 3241

Qy         1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3242 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3301

Qy         1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3302 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3361

Qy         1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3362 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3421

Qy         1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3422 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3481

Qy         1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3482 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3541

Qy         1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3542 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3601

Qy         1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3602 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3661

Qy         1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3662 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3721

Qy         1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3722 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3781

Qy         1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3782 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3841

Qy         1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3842 GAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCAGGACCC 3901

Qy         1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3902 CCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTGGGCAGA 3961

Qy         1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3962 GCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4021

Qy         1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4022 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGGGATCCCT 4081

Qy         1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4082 CGGTCCTACGTGGAGGCACAGGAGGACTACCTGCGTCGGGAGGCCAAGCTCCTAAAGCGG 4141

Qy         1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4142 GAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGACCTGACCGAGGCCTACAAGACGCAG 4201

Qy         1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTGAAGGAG 4261

Qy         1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4262 GCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAGCTGCCC 4321

Qy         1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4322 CTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAGTACGAC 4381

Qy         1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4382 ACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4441

Qy         1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4442 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4501

Qy         1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4502 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4561

Qy         1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4562 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGTAAGCCGCGGCAGAGCCCC 4621

Qy         1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4622 CTGACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4681

Qy         1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4682 ACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAGGCATCC 4741

Qy         1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4742 CAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGCACCGTG 4801

Qy         1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4802 CCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTGAGTGGC 4861

Qy         1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4862 GTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCCCGCGGC 4921

Qy         1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4922 ATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4981

Qy         1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4982 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5041

Qy         1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5042 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACACGGCC 5101

Qy         1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5102 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5161

Qy         1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5162 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5221

Qy         1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5222 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5281

Qy         1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5282 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5341

Qy         1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5342 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5401

Qy         1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5402 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5461

Qy         1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         5462 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5521
```

```
Qy        1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5522 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5581

Qy        1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5582 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5641

Qy        1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5642 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCAAGCCCACGGTCCTGAGGTCC 5701

Qy        1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5702 ACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5761

Qy        1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5762 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5821

Qy        1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5822 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5881

Qy        1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5882 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5941

Qy        1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5942 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6001

Qy        2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6002 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6061

Qy        2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6062 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6121

Qy        2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6122 CTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 6181

Qy        2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6182 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 6241

Qy        2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6242 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 6301

Qy        2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6302 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6361

Qy        2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6362 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6421

Qy        2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6422 GAGGTCATCACACAGGACTACACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCC 6481

Qy        2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6482 GCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCC 6541

Qy        2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6542 AGTGACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGC 6601

Qy        2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6602 GAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6661

Qy        2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6662 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6721

Qy        2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6722 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6781

Qy        2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6782 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6841

Qy        2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6842 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6901

Qy        2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6902 GAGCCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACC 6961

Qy        2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6962 GGAACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATG 7021

Qy        2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7022 GGGCTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCC 7081

Qy        2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7082 CCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCT 7141

Qy        2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7142 ATGCCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGC 7201

Qy        2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7202 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7261

Qy        2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7262 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7321

Qy        2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThr 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7322 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACG 7381

Qy        2461 ProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProPro 2480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7382 CCATTCCCCTACAACCCCCTGATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCC 7441

Qy        2481 ProProGlyLeuProAlaGlySerGlyProLeuAlaGlyProHisHisAlaTrpAspGlu 2500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7442 CCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAG 7501

Qy        2501 GluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7502 GAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAG 7552

RESULT 2
AF125672
LOCUS       AF125672                8686 bp    mRNA    linear   PRI 04-APR-1999
DEFINITION  Homo sapiens silencing mediator of retinoic acid and thyroid
            hormone receptor extended isoform (SMRTE) mRNA, complete cds.
ACCESSION   AF125672
VERSION     AF125672.1  GI:4559297
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

AAD 22973

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8686)
  AUTHORS   Park,E.J., Schroen,D.J., Yang,M., Li,H., Li,L. and Chen,J.D.
  TITLE     SMRTe, a silencing mediator for retinoid and thyroid hormone
            receptors-extended isoform that is more related to the nuclear
            receptor corepressor
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)
  MEDLINE   99199215
   PUBMED   10097068
REFERENCE   2  (bases 1 to 8686)
  AUTHORS   Chen,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-1999) Pharmacology and Molecular Toxicology,
            University of Massachusetts Medical School, 55 Lake Avenue North,
            Worcester, MA 01655, USA
FEATURES             Location/Qualifiers
     source          1. .8686
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .8686
                     /gene="SMRTE"
     CDS             157. .7680
                     /gene="SMRTE"
                     /function="transcriptional corepressor"
                     /note="SMRTe"
                     /codon_start=1
                     /product="silencing mediator of retinoic acid and thyroid
                     hormone receptor extended isoform"
                     /protein_id="AAD22973.1"
                     /db_xref="GI:4559298"
                     /translation="MSGSTQPVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQH
                     HSRDYASHLSPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEF
                     IESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPEL
                     ELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPP
                     IESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAM
                     RKKLILYFKRRNHARKQWEQKFCQRYDQLMEAWEKKVERIENNPRRRAKESKVREYYE
                     KQFPEIRKQRELQERMQRVGQRGSGLSMSPARSEHEVSEIIDGLSEQENLEKQMRQLA
                     VIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNF
                     GLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQ
                     QPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASK
                     GRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETAKK
                     GLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKK
                     KKAPAAASEEAAFPPVVEDEEMEASGVTGNEEEMVEEAEATVNNSSDTESIPSPHTEA
                     AKDTGQNGPKPPATLGADGPPPGPPTPPPEDIPAPTESTPASEATLAPTPPPAPPFPS
                     SPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGP
                     AKGKDAEAAEATAERALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGG
                     DKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAA
                     PTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAEKPVFFPAFAAEAQ
                     KLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR
                     PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPM
                     DPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPS
                     TRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEG
                     KKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGL
                     MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPP
                     PSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPL
                     KEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACY
                     EESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTM
                     REPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSG
                     VDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAA
                     LENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQ
                     VPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSS
                     ERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPA
                     SHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPA
                     ATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSG
                     LEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKT
                     QSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEG
                     ELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEV
                     ITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHS
                     EGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMG
                     SKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNM
                     PAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNAS
```

```
                     ASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSS
                     VHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGP
                     LAGPHHAWDEEPKPLLCSQYETLSDSE"
ORIGIN

Alignment Scores:
Pred. No.:              7.81e-254       Length:          8686
Score:                  12978.00         Matches:         2487
Percent Similarity:     98.57%          Conservative:    3
Best Local Similarity:  98.46%          Mismatches:      9
Query Match:            98.21%          Indels:          28
DB:                     9               Gaps:            3

US-09-522-753-5 (1-2517) x AF125672 (1-8686)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        157 ATGTCGGGCTCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 216

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 276

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 336

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 396

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 456

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 516

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 576

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 636

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 696

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 756

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 816

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 876

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 936

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 996

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db           997 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 1056

Qy           301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                 ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db          1057 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1116

Qy           321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1117 CGGCGGCGGGCCAAGGAGAGCAAGGTTCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1176

Qy           341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                 |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db          1177 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1233

Qy           361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                 ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db          1234 CTGTCCATGTCGCCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1293

Qy           381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1294 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1353

Qy           401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1354 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1413

Qy           421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1414 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1473

Qy           441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1474 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1533

Qy           461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1534 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1593

Qy           481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1594 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACAGCAGCAG 1653

Qy           501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1654 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAGAAA 1713

Qy           521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1714 GATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAGAAC 1773

Qy           541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1774 GACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1833

Qy           561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1834 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1893

Qy           581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1894 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1953

Qy           601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1954 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 2013

Qy           621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2014 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 2073

Qy           641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2074 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2133

Qy           661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2134 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2193

Qy           681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2194 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2253

Qy           701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                 ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db          2254 GAGGATGAGGAGATGGAGGCGTCGGGCGTGACGGGAAATGAGGAGGAGATGGTGGAGGAG 2313

Qy           721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                 ||||||                                                   |||
Db          2314 GCTGAA---------------------------------------------------GCC 2322

Qy           741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2323 ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGCCAAG 2382

Qy           761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2383 GACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGCCACCCCCA 2442

Qy           781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrProAlaSe 800
                 |||||||||||||||||| |||||||||||||||   |||   |||   |||||||||||
Db          2443 GGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCCTC 2501

Qy           800 rGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPr 820
                 ||||||||||   |||||||||||||||||||||||||||   ||||||:::||||||||
Db          2502 TGAAGCCACCTTAGCCCCTACGCCCCCACCAGCACCCCCATTTCCCTCTTCACCTCCTCC 2561

Qy           820 oValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGl 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2562 TGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGG 2621

Qy           840 yGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGl 860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2622 GGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGA 2681

Qy           860 uProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAl 880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2682 GCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGC 2741

Qy           880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySe 900
                 |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db          2742 GGAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAG 2801

Qy           900 rGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaTh 920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2802 CGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCAC 2861

Qy           920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPr 940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2862 CTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCC 2921

Qy           940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPr 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2922 AAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCC 2981

Qy           960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLy 980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2982 ACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAA 3041

Qy           980 sValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3042 AGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACC 3101

Qy          1000 oProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProAr 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3102 GCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCG 3161
```

```
Qy   1020 gGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------------- 1033
          ||||||||||||||||||||||||||||||||||||||||
Db   3162 GGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTT 3221

Qy   1034 ----AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGl 1052
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3222 CCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGG 3281

Qy   1052 yLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPr 1072
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3282 CCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCC 3341

Qy   1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3342 CTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGACACTGC 3401

Qy   1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAl 1112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3402 CCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGC 3461

Qy   1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGl 1132
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3462 CAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCA 3521

Qy   1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3522 GCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCT 3581

Qy   1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLe 1172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3582 GCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCT 3641

Qy   1172 uSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAl 1192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3642 GTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGC 3701

Qy   1192 aSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIl 1212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3702 GTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCAT 3761

Qy   1212 eProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGl 1232
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3762 TCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCACCCACGG 3821

Qy   1232 yThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerPr 1252
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3822 CACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCC 3881

Qy   1252 oSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGl 1272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3882 GAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGG 3941

Qy   1272 yLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGl 1292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3942 CAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGA 4001

Qy   1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4002 GGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGA 4061

Qy   1312 pMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGl 1332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4062 CATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCTCATGGG 4121

Qy   1332 yArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHisIleAr 1352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4122 CCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCACATCCG 4181

Qy   1352 gGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuAr 1372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4182 CGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTACCTGCG 4241

Qy   1372 gArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArgAs 1392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4242 TCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTCACGGGA 4301

Qy   1392 pLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGl 1412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4302 CCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGA 4361

Qy   1412 uGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGl 1432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4362 GGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGA 4421

Qy   1432 uLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleTh 1452
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4422 GCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCAC 4481

Qy   1452 rGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAs 1472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGA 4541

Qy   1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4542 CGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGT 4601

Qy   1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgPr 1512
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4602 GATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCC 4661

Qy   1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4662 AGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGA 4721

Qy   1532 uLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGl 1552
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4722 GCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTTTGCCGG 4781

Qy   1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGl 1572
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4782 CCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCAGGAGGG 4841

Qy   1572 ySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIl 1592
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4842 CAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGAT 4901

Qy   1592 eAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGl 1612
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4902 CGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGA 4961

Qy   1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPh 1632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4962 GCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTT 5021

Qy   1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuPr 1652
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5022 CGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCC 5081

Qy   1652 oArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGl 1672
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5082 CCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGG 5141

Qy   1672 yTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSe 1692
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5142 CTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTC 5201

Qy   1692 rGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5202 GCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGG 5261

Qy   1712 yLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIl 1732
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5262 CCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCAT 5321

Qy   1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5322 CGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGC 5381

Qy        1752 aThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHi 1772
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5382 CACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCA 5441

Qy        1772 sSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSe 1792
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5442 CAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTC 5501

Qy        1792 rSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLy 1812
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5502 CTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAA 5561

Qy        1812 sSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGl 1832
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5562 GTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGA 5621

Qy        1832 uGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAl 1852
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5622 GCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGC 5681

Qy        1852 aSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5682 CTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCA 5741

Qy        1872 nGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPr 1892
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5742 GCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCC 5801

Qy        1892 oSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaTh 1912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5802 CAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCCAC 5861

Qy        1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5862 ATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCT 5921

Qy        1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5922 CATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCG 5981

Qy        1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5982 AGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGC 6041

Qy        1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6042 CTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGC 6101

Qy        1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6102 CACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCC 6161

Qy        2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6162 GGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTC 6221

Qy        2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6222 CATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGG 6281

Qy        2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6282 GGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAA 6341

Qy        2072 sHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGl 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6342 GCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGG 6401

Qy        2092 yProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSe 2112
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6402 CCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAG 6461
```

```
Qy        2112 rGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVa 2132
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6462 CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGT 6521

Qy        2132 lValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPr 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6522 GGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCACCC 6581

Qy        2152 oGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPr 2172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6582 ACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGCCC 6641

Qy        2172 oValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAl 2192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6642 CGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGTGC 6701

Qy        2192 aProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysTh 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6702 CCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGAC 6761

Qy        2212 rSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetTh 2232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6762 GTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATGAC 6821

Qy        2232 rGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnTh 2252
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6822 GGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAGAC 6881

Qy        2252 rGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePh 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6882 GGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTT 6941

Qy        2272 eSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLy 2292
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAA 7001

Qy        2292 sLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGl 2312
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7002 GAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACGGA 7061

Qy        2312 uIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVa 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7062 GATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCGGT 7121

Qy        2332 lGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGl 2352
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7122 GCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATGGG 7181

Qy        2352 yLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAs 2372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7182 TAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTGAA 7241

Qy        2372 nAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHi 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7242 TGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGACCA 7301

Qy        2392 sThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerAr 2412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7302 CACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCG 7361

Qy        2412 gLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSerSe 2432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7362 AAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTC 7421

Qy        2432 rValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAs 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7422 AGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGA 7481

Qy        2452 pArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGl 2472
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7482 CAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTGCA 7541
```

```
Qy        2472 nAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeuAl 2492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7542 GGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTCGC 7601

Qy        2492 aGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluTh 2512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7602 TGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAGAC 7661

Qy        2512 rLeuSerAspSerGlu 2517
               ||||||||||||||||
Db        7662 ACTCTCCGACAGCGAG 7677

RESULT 3
AR447713
LOCUS       AR447713                9053 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 306 from patent US 6673549.
ACCESSION   AR447713
VERSION     AR447713.1  GI:42676037
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 9053)
  AUTHORS   Furness,L.M. and Buchbinder,J.L.
  TITLE     Genes expressed in C3A liver cell cultures treated with steroids
  JOURNAL   Patent: US 6673549-A 306 06-JAN-2004;
FEATURES             Location/Qualifiers
     source          1. .9053
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.09e-252      Length:         9053
Score:                  12922.00        Matches:        2487
Percent Similarity:     98.26%         Conservative:   2
Best Local Similarity:  98.18%         Mismatches:     10
Query Match:            97.78%         Indels:         35
DB:                     6              Gaps:           5

US-09-522-753-5 (1-2517) x AR447713 (1-9053)

Qy           1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db         520 ATGTCGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 579

Qy          21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         580 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 639

Qy          41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         640 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 699

Qy          61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         700 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 759

Qy          81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         760 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 819

Qy         101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         820 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 879

Qy         121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         880 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 939

Qy         141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         940 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 999

Qy         161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1000 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 1059

Qy         181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1060 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 1119

Qy         201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1120 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 1179

Qy         221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1180 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 1239

Qy         241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1240 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 1299

Qy         261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1300 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 1359

Qy         281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
               ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        1360 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1419

Qy         301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
               ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        1420 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1479

Qy         321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1480 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1539

Qy         341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1540 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1599

Qy         361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1600 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1659

Qy         381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
               ||||||||||||   :::||||||||||||||||||||||||||||||||||||||||||
Db        1660 GAGCAGGAGAACCGGAAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1719

Qy         401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1720 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1779

Qy         421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1780 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1839

Qy         441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1840 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1899

Qy         461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1900 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1959

Qy         481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGln---GlnGlnGlnGln 499
               |||||||||||||||||||||||||||||||||||||||||||||   |||   ||||||
Db        1960 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCNACAACNGCAGCAG 2019

Qy         500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2020 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAG 2079
```

```
Qy          520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2080 AAAGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAG 2139

Qy          540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2140 AACGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGAC 2199

Qy          560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2200 GAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAA 2259

Qy          580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2260 GGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAG 2319

Qy          600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2320 CAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAA 2379

Qy          620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2380 ATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGG 2439

Qy          640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2440 ATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2499

Qy          660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2500 CAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCG 2559

Qy          680 ArgArgLysLysLysLysAlaProAlaAlaAla-SerGluGluAlaAlaPheProProVa 699
                ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         2560 CGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCCAGCGAGGAGGCTGCATTCCCGCCCGT 2619

Qy          699 lValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGl 719
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2620 GGTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGA 2679

Qy          719 uGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPr 739
                ||||||||||
Db         2680 GGAGGCTGAA-------------------------------------------------- 2689

Qy          739 oAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAl 759
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2690 -GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGC 2748

Qy          759 aLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAsp-GlyProP 779
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         2749 CAAGGACACAGGGCAGAATGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGGCCAC 2808

Qy          779 roProGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrPro 798
                ||||||||||||||||||||||| ||||||||||||||||  |||   ||||||||||||
Db         2809 CCCCAGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCG 2867

Qy          799 AlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaPro 818
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2868 GCCTCTGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCT 2927

Qy          819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGlu 838
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2928 CCTCCTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAG 2987

Qy          839 GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2988 GAGGGGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCC 3047

Qy          859 GluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLys 878
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3048 GAGGAGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAG 3107

Qy          879 AspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGly 898
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3108 GACGCGGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGC 3167

Qy          899 GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3168 GGGAGCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGT 3227

Qy          919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3228 GCTACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTG 3287

Qy          939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3288 TCCCCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAG 3347

Qy          959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         3348 AAGCCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATC---GTC 3404

Qy          979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProPro-Al 998
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3405 ACCAAAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGAC 3464

Qy          998 aProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSe 1018
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3465 CCCACCGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAG 3524

Qy         1018 rProArgGlyLysSerArgSerProAlaProProAlaAspLysGlu-------------- 1033
                ||||||||||||||||||||||||||||||||||||||||||||||
Db         3525 CCCCCGGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGT 3584

Qy         1034 ----------AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpTh 1050
                          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3585 GTTCTTCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGAC 3644

Qy         1050 rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3645 TTCCGGCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCC 3704

Qy         1070 oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3705 GGACCCCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGA 3764

Qy         1090 pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3765 CACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTC 3824

Qy         1110 rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3825 CTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTC 3884

Qy         1130 rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3885 GGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCAT 3944

Qy         1150 tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3945 GGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGA 4004

Qy         1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4005 GCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCA 4064

Qy         1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4065 GGAGGCGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAA 4124

Qy         1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4125 AGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCAC 4184

Qy         1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4185 CCACGGCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGA 4244

Qy     1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4245 CAGCCCGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTA 4304

Qy     1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4305 CGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTC 4364

Qy     1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4365 CAAGGAGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCAC 4424

Qy     1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4425 CTATGACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCT 4484

Qy     1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4485 CATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCA 4544

Qy     1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4545 CATCCGCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTA 4604

Qy     1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4605 CCTGCGTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTC 4664

Qy     1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4665 ACGGGACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGC 4724

Qy     1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4725 CCATGAGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCG 4784

Qy     1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySe 1450
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4785 CGAGGAGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTC 4844

Qy     1450 rIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLy 1470
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4845 CATCACGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAA 4904

Qy     1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLe 1490
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4905 GCACGACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCT 4964

Qy     1490 uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSe 1510
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4965 GGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAG 5024

Qy     1510 rArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa 1530
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5025 CCGGCCAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGT 5084

Qy     1530 lProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPh 1550
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5085 GCCTGAGCTGGGCAAGCCGCGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTT 5144

Qy     1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl 1570
            ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db     5145 TGCCGGCCACCTCCCACGAGGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCA 5204

Qy     1570 nGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5205 GGAGGGCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCG 5264

Qy     1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5265 TGAGATCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCC 5324
```

```
Qy     1610 oTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLe 1630
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5325 CTATGAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCT 5384

Qy     1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTy 1650
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5385 GGCCTTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTA 5444

Qy     1650 rLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIl 1670
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5445 CCTGCCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCAT 5504

Qy     1670 eArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIl 1690
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5505 CCGCGGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACAT 5564

Qy     1690 eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
            |||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db     5565 CACCTCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCT 5624

Qy     1710 uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGl 1730
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5625 GAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGG 5684

Qy     1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5685 CATCATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCAC 5744

Qy     1750 rProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSe 1770
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5745 CCCAGCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAG 5804

Qy     1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh 1790
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5805 CCGCCACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCAC 5864

Qy     1790 rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr 1810
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5865 CACGTCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCG 5924

Qy     1810 gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGl 1830
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5925 GGAAAAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGG 5984

Qy     1830 yThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerAr 1850
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     5985 TACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCG 6044

Qy     1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6045 CCCCGCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGC 6104

Qy     1870 aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa 1890
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6105 CCTCCAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGT 6164

Qy     1890 lGluProSerLysProThrValLeu---------ArgSerThrSerThrSerSerProVa 1907
            ||||||||||   ||||||||||||         ||||||||||||||||||||||||||
Db     6165 GGAGCCCAGCACGCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCCTCACCCGT 6224

Qy     1907 lArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGl 1927
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6225 TCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGG 6284

Qy     1927 yValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaAr 1947
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6285 GGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCG 6344

Qy     1947 gProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSe 1967
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6345 GCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTC 6404
```

```
Qy        1967 rGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProPr 1987
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6405 CGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCC 6464

Qy        1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAl 2007
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6465 TGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGC 6524

Qy        2007 aSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGl 2027
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6525 CAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCA 6584

Qy        2027 nSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSe 2047
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6585 AAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAG 6644

Qy        2047 rTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAs 2067
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6645 CTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGA 6704

Qy        2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuAr 2087
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6705 CAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCG 6764

Qy        2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6765 GCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCG 6824

Qy        2107 gProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLy 2127
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6825 GCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAA 6884

Qy        2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6885 AGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTA 6944

Qy        2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6945 CACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCC 7004

Qy        2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7005 TGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCC 7064

Qy        2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerPr 2207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7065 CCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCC 7124

Qy        2207 oGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerPr 2227
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7125 AGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCC 7184

Qy        2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7185 ACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCG 7244

Qy        2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGl 2267
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7245 GGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCA 7304

Qy        2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLy 2287
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7305 GCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAA 7364

Qy        2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7365 GCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAG 7424

Qy        2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7425 CCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTA 7484

Qy        2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7485 TAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAG 7544

Qy        2347 gLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAl 2367
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7545 AAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGC 7604

Qy        2367 aPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAs 2387
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7605 TTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGA 7664

Qy        2387 pGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGl 2407
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7665 CGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGG 7724

Qy        2407 yArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPr 2427
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7725 CAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCC 7784

Qy        2427 oProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAs 2447
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7785 ACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAA 7844

Qy        2447 nArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLe 2467
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7845 CCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCT 7904

Qy        2467 uIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGl 2487
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7905 GATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGG 7964

Qy        2487 ySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCy 2507
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        7965 CAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTG 8024

Qy        2507 sSerGlnTyrGluThrLeuSerAspSerGlu 2517
               ||||||||||||||||||||||||||||||
Db        8025 CTCGCAGTACGAGACACTCTCCGACAGCGAG 8055

RESULT 4
AF113001
LOCUS       AF113001                8388 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus silencing mediator of retinoic acid and thyroid
            hormone receptor alpha mRNA, complete cds.
ACCESSION   AF113001
VERSION     AF113001.1  GI:4454547
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 8388)
  AUTHORS   Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and
            Evans,R.M.
  TITLE     Unique forms of human and mouse nuclear receptor corepressor SMRT
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)
  MEDLINE   99178941
   PUBMED   10077563
REFERENCE   2  (bases 1 to 8388)
  AUTHORS   Downes,M.R., Ordentlich,P. and Evans,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk
            Institute for Biological Studies, 10010 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1..8388
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="spleen; brain"
     CDS             626..8047
                     /function="transcriptional co-repressor"
                     /note="mSMRT alpha"
                     /codon_start=1
                     /product="silencing mediator of retinoic acid and thyroid
                     hormone receptor alpha"
                     /protein_id="AAD20944.1"
                     /db_xref="GI:4454548"
                     /translation="MSGSTQPVAQTWRAAEPRYPPHGISYPVQIARSHTDVGLLEYQH
                     HPRDYTSHLSPGSIIQPQRRRPSLLSEFQPGSERSQELHLRPESRTFLPELGKPDIEF
                     TESKRPRLELLPDTLLRPSPLLATGQPSGSEDLTKDRSLAGKLEPVSPPSPPHADPEL
                     ELAPSRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPP
                     IESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAM
                     RKKLILYFKRRNHARKQWEQRFCQRYDQLMEAWEKKVERIENNPRRRAKESKVREYYE
                     KQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQL
                     AVIPPMLYDADQQRIKFINMNGLMDDPMKVYKDRQVTNMWSEQERDTFREKFMQHPKN
                     FGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQM
                     ARSSQEEKEEKEKEKEADKEEEKQDAENEKEELSKEKTDDTSGEDNHEKEAVASKGRK
                     TANSQGRRKGRITRSMANEANHEETATPQQSSELASMEMNESSRWTEEEMETAKKGLL
                     EHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKT
                     PAAASEETAFPPAAEDEEMEASGASANEEELAEEAEASQASGNEVPRVGECSGPAAVN
                     NSSDTESVPSPRSEAMKDTGPKPTGTEALPAATQPPVPPPEEPAVAPAEPSPVPDASG
                     PPSPEPSHHLPHPRLLWTRMNKKPRLLQLPRQRMPRSRSLRPRRSMWEKPEEPEASEE
                     PPESVKSDHKEETEEEPEDKAKGTEAIETVSEAPLKVEEAGSKAAVTKGSSSGATQDS
                     DFSATCSADEVDEPEGGDKGRLLSPRPSLLTPAGDPRASTSPQKPLDLKQLKQRAAAI
                     PPIQVTKVHEPPREDTVPPKPVPPVPPPTQHLQPEGDVSQQSGGSPRGKSRSPVPPAE
                     KEAEKPAFFPAFPTEGPKLPTEPPRWSSGLPFPIPPREVIKTSPHAADPSAFSYTPPG
                     HPLPLGLHDSARPVLPRPPISNPPPLISSAKHPGVLERQLGAISQQGMSVQLRVPHSE
                     HAKAPMGPLTMGLPLAVDPKKLGTALGSATSGSITKGLPSTRAADGPSYRGSITHGTP
                     ADVLYKGTISRIVGEDSPSRLDRAREDTLPKGHVIYEGKKGHVLSYEGGMSVSQCSKE
                     DGRSSSGPPHETAAPKRTYDMMEGRVGRTVTSASIEGLMGRAIPEQHSPHLKEQHHIR
                     GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPRDLTETYKPRPLDPLGPLKL
                     KPTHEGVVATVKEAGRSIHEIPREELRRTPELPLAPRPLKEGSITQGTPLKYDSGAPS
                     TGTKKHDVRSIIGSPGRPFPALHPLDIMADARALERACYEESLKSRSGTSSGAGGSIT
                     RGAPVVVPELGKPRQSPLTYEDHGAPFTSHLPRGSPVTTREPTPRLQEGSLLSSKASQ
                     DRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVTGVDLYRGHIPLAFDPTSIPR
                     GIPLEAAAAAYYLPRHLAPSPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMH
                     HNAASAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATA
                     IDRLAYLPTAPPPFSSRHSSSPLSPGGPTHLAKPTATSSSERERERERERDKSILTST
                     TTVEHAPIWRPGTEQSSGAGGSSRPASHTHQHSPISPRTQDALQQRPSVLHNTSMKGV
                     VTSVEPGTPTVLRWARSTSTSSPVRPAATFPPATHCPLGGTLEGVYPTLMEPVLLPKE
                     TSRVARPERARVDAGHAFLTKPPGREPASSPSKSSEPRSLAPPSSSHTAIARTPAKNL
                     APHHASPDPPAPTSASDLHREKTQSKPFSIQELELRSLGYHSGAGYSPDGVEPISPVS
                     SPSLTHDKGLSKPLEELEKSHLEGELRHKQPGPMKLSAEAAHLPHLRPLPESQPSSSP
                     LLQTAPGIKGHQRVVTLAQHISEVITQDYTRHHPQQLSGPLPAPLYSFPGASCPVLDL
                     RRPPSDLYLPPPDHGTPARGSPHSEGGKRSPEPSKTSVLGSSEDAIEPVSPPEGMTEP
                     GHARSTAYPLLYRDGEQGEPRMGLESPGNTSQPPTFFSKLTESNSAMVKSKKQEINKK
                     LNTHNRNEPEYNIGQPGTEIFNMPAITGAGLMTCRSQAVQEHASTNMGLEAIIRKALM
                     GKYDQWEEPPPLGANAFNPLNASASLPAAAMPITTADGRSDHALTSPGGGGKAKVSGR
                     PSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNP
                     LIMRLQAGVMASPPPPGLAAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE"
     misc_feature    731. .1387
                     /note="Region: deleted in SMRT beta isoform encoded by
                     GenBank Accession Number AF113002"
ORIGIN

Alignment Scores:
Pred. No.:              2.74e-215     Length:         8388
Score:                  11067.00       Matches:        2160
Percent Similarity:     88.20%        Conservative:   89
Best Local Similarity:  84.71%        Mismatches:     192
Query Match:            83.75%        Indels:         110
DB:                     10            Gaps:           29

US-09-522-753-5 (1-2517) x AF113001 (1-8388)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db        626 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 685

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db        686 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 745

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db        746 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 805

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        806 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGGTCTCAG 865

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||   ::::::||||||||||||||||||   ::::::
Db        866 GAGCTCCACCTGCGCCCTGAGTCCCGCACGTTCCTGCCTGAGCTGGGCAAGCCCGACATA 925

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||   ||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        926 GAATTCACCGAGAGCAAGCGCCCCCGCCTGGAGCTACTACCCGATACCCTGCTGCGCCCA 985

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        986 TCACCCCTGCTGGCCACTGGGCAGCCGAGTGGGTCTGAAGACCTTACCAAGGACCGTAGC 1045

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db       1046 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 1105

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db       1106 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 1165

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 CGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAGTTG 1225

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 GAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATAGAA 1285

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 TCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGAAGAAAGCCGAAGCC 1345

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 GCACACCGGATCCTAGAAGGCCTGGGGCCCCAGGTGGAGCTGCCTCTGTACAACCAGCCG 1405

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 TCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTG 1465

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||:::|||||||||
Db       1466 ATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGAACCAGCGCTTCTGCCAG 1525

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1526 CGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAATCCG 1585

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAGATC 1645

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 CGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGTGGG 1705

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 CTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTGTCT 1765

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGTTGTAC 1825

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db       1826 GACGCGGACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCCATG 1885
```

```
Qy         421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
               ||||||||||||||||||||||||   |||||||||||||||||||||::::::||||||
Db        1886 AAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACCTTC 1945

Qy         441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1946 CGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAGAGA 2005

Qy         461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2006 AAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTACAAG 2065

Qy         481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2066 AGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAGCAA 2125

Qy         501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
               |||||||||||||||||||||            |||   |||||||||||||||||||||
Db        2126 CAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAGAAG 2173

Qy         521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
               :::||||||||||||||||||||||||:::|||||||||||||||   :::   ||||||
Db        2174 GAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAGAAC 2233

Qy         541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
               :::||||||:::|||   ||||||||||||||||||||||||||||||||||||   |||
Db        2234 GAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACCATGAG 2293

Qy         561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2294 AAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAAGGC 2353

Qy         581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
               |||||||||||||||||||||||||||||||||   ||||||      ||||||||||||
Db        2354 CGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAGCAA 2413

Qy         601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
               |||:::||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        2414 AGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAGATG 2473

Qy         621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2474 GAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGCATG 2533

Qy         641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2534 GTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 2593

Qy         661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2594 AACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCTCGG 2653

Qy         681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
               |||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db        2654 AGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCCGCT 2713

Qy         701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
               |||||||||||||||||||||||||||   |||   ||||||||||||:::   ||||||
Db        2714 GAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAGGAG 2773

Qy         721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGlyPro 739
               |||||||||      ||||||||||||||||||||||||   ||||||||||||||||||
Db        2774 GCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGCCCA 2833

Qy         740 AlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAla 759
               |||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db        2834 GCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCCATG 2893

Qy         760 LysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAspGly 777
               |||||||||         ||||||||||||      |||      |||   |||
Db        2894 AAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCCACC 2944

Qy         778 ProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluProThr 797
                  ||||||   |||      ||||||         :::   ||||||   ||||||:::
Db        2945 CAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGTAGCCCCTGCTGAGCCCTCC 2998

Qy         798 ProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAla 817
               |||      :::|||:::|||   |||:::|||   |||      |||||||||:::|||
Db        2999 CCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTC----CCATCACCTGCCGCA 3054

Qy         818 ProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProVal 837
               ||||||      |||   |||:::|||:::|||         |||   ||||||
Db        3055 CCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAGACA 3114

Qy         838 GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLys 857
               |||:::   :::|||||||||      |||   |||         :::|||   ||||||
Db        3115 GAGGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGGAAAA 3165

Qy         858 Ala-GluGluPro---------------------------ValLysSerGluCysThrGl 868
               ||| |||||||||                           |||||||||:::      ||
Db        3166 GCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTAAAGAGTGACCACAAGGA 3225

Qy         868 uGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThrAl 887
               ||||   ||||||   |||      |||   |||      ||||||   |||      ::
Db        3226 GGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTGTC 3285

Qy         887 aGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThr---Al 906
               :|||      ||||||   |||      |||   ||||||      |||   |||
Db        3286 TGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAGGCAGCTGTGACCAAGGG 3339

Qy         906 aLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluVa 926
                   ||||||||||||   ||||||||||||   ||||||||||||||||||||||||||
Db        3340 TTCCAGCTCAGGTGCCACCCAGGACAGTGACTTCAGTGCCACCTGCAGTGCCGATGAGGT 3399

Qy         926 lAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuTh 946
               |||||||   |||||||||||||||   ||||||||||||||||||||||||||||||||
Db        3400 GGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTCAC 3459

Qy         946 rProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLe 966
               ||||   |||||||||||||||:::   ||||||||||||||||||||||||||||||||
Db        3460 CCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAGCT 3519

Qy         966 uLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProAr 986
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3520 GAAGCAGCGAGCAGCCGCCATCCCCCCTATCCAGGTCACCAAGGTCCATGAGCCCCCCCG 3579

Qy         986 gGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGl 1006
               |||||||      |||   ||||||   ||||||   |||||||||   |||:::|||||
Db        3580 GGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTACA 3639

Qy        1006 nProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerPr 1026
               |||||||   |||      ||||||   |||   ||||||||||||||||||||||||||
Db        3640 GCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGCCC 3699

Qy        1026 oAlaProProAlaAspLysGlu------------------------AlaPheAlaAlaGl 1038
               |   |||||||||:::||||||                        ||||||      ||
Db        3700 AGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACTGA 3759

Qy        1038 uAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValPr 1058
               |      |||||||||   :::||||||   |||:::||||||||||||||||||:::||
Db        3760 GGGCCCAAAGCTACCGACTGAGCCCCCACGCTGGTCATCGGGCCTGCCCTTCCCCATCCC 3819

Qy        1058 oProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAl 1078
               |||||||||||||||||||   ||||||||||||   |||||||||||||||||||||
Db        3820 TCCACGGGAGGTGATCAAGACTTCCCCACACGCCGCTGACCCCTCTGCCTTCTCCTACAC 3879

Qy        1078 aProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProAr 1098
                ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        3880 ACCCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCACG 3939

Qy        1098 gProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLe 1118
               |||||||   ||||||||||||||||||||||||||||||||||||||||||   |||||
Db        3940 TCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTACT 3996

Qy        1118 uGluArgGlnIleGlyAlaIleSer---GlnGlyMetSerValGlnLeuHisValProTy 1137
```

```
              ||||||||||:::||||||||||||   |||||||||||||||||||||   ||||||::
Db       3997 TGAGAGGCAGCTGGGTGCCATCTCCCAGCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCA 4056

Qy       1137 rSerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAs 1157
              :|||||||||||||||||||||:::||||||:::||||||||||||||||||   :::||
Db       4057 CTCAGAGCATGCCAAGGCCCCCATGGGCCCTCTCACCATGGGGCTGCCCCTTGCCGTGGA 4116

Qy       1157 pProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGl 1177
              |||||||||||||
Db       4117 CCCTAAGAAGCTG----------------------------------------------- 4129

Qy       1177 nAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGl 1197
                                                                        ||
Db       4130 ----------------------------------------------------------GG 4131

Qy       1197 yThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVa 1217
              ||||||||||||||||         ||||||||||||||||||:::||||||||||||
Db       4132 GACAGCACTGGGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCTCCCCAGTACCCGGGC 4191

Qy       1217 lProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVa 1237
                                 :::||||||||||||||||||||||||||||||||||||||
Db       4192 TGCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGGCACGCCCGCAGACGT 4245

Qy       1237 lLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspAr 1257
              |||||||||||||||||||:::||||||:::|||||||||||||||||||||||||||||
Db       4246 CCTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCG 4305

Qy       1257 gGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisVa 1277
              |   |||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       4306 GGCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGT 4365

Qy       1277 lLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSe 1297
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db       4366 CCTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAG 4425

Qy       1297 rSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyAr 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4426 CTCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCG 4485

Qy       1317 gValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPr 1337
              ||||||||||   ::::::||||||||||||||||||||||||||||||||||||   ||
Db       4486 TGTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CC 4542

Qy       1337 oGluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGl 1357
              ||||:::|||||||||   |||||||||||||||||||||||||||||||||||||||||
Db       4543 TGAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCA 4599

Qy       1357 nGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLe 1377
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4600 AGGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGAGGCCAAGCT 4659

Qy       1377 uLeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTy 1397
              |||||||||||||||||||||||||||||||||||||   |||||||||||||||   ||
Db       4660 CTTGAAGCGAGAAGGGACACCACCACCCCCACCACCACCTCGGGACCTGACTGAGACCTA 4719

Qy       1397 rLysThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLe 1414
              ||||   :::            ||||||||||||||||||||||||   |||||||||::
Db       4720 CAAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGT 4779

Qy       1414 uValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuAr 1434
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4780 GGTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCG 4839

Qy       1434 gHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGl 1454
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4840 CCGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGG 4899

Qy       1454 yThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValAr 1474
              |||||||||||||||||||:::||||||   :::||||||:::|||||||||||||||||
Db       4900 CACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCG 4959

Qy       1474 gSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAl 1494
              ||||:::||||||||||||||||||   ||||||   :::||||||||||||:::|||||
Db       4960 CTCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGC 5019

Qy       1494 aAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyTh 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db       5020 TGACGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGAC 5079

Qy       1514 rAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGl 1534
              |:::|||   :::||||||||||||   |||||||||||||||:::||||||||||||||
Db       5080 CAGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGG 5139

Qy       1534 yLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLe 1554
              |||||||||||||||||||||||||||||||||||||||||||||||||      |||||
Db       5140 CAAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCT 5199

Qy       1554 uProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLe 1574
              ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db       5200 GCCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCT 5259

Qy       1574 uSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLy 1594
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5260 CCTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAA 5319

Qy       1594 sSerProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLe 1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5320 GTCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTT 5379

Qy       1614 uLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPr 1634
              |||||||||||||:::||||||||||||||||||   |||||||||||||||||||||||
Db       5380 GCTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCC 5439

Qy       1634 oThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProAr 1653
              ||||||||||||||||||||||||||||:::   ||||||||||||||||||||||||||
Db       5440 CACCTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCG 5499

Qy       1653 gHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTy 1673
              |||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db       5500 GCACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTA 5559

Qy       1673 rProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGl 1693
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5560 CCCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCA 5619

Qy       1693 nGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLe 1713
              ||||||||||||||||   |||:::|||||||||||||||||||||||||||||||||||
Db       5620 GCAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCT 5679

Qy       1713 uSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAs 1733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5680 GTCACCGCGAGAGTCCTCGCTGGCCCTCAATTATGCCGCTGGCCCAAGAGGCATTATCGA 5739

Qy       1733 pLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaTh 1753
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5740 CCTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCAC 5799

Qy       1753 rAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSe 1773
              ||||:::||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       5800 CGCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAG 5859

Qy       1773 rSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSe 1793
              |||||||||||||||||||||||||||||||||||||   |||||||||   ||||||||
Db       5860 TAGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTC 5919

Qy       1793 rSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSe 1813
              ||||||||||||||||:::|||:::||||||||||||                  |||||
Db       5920 ATCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTC 5961

Qy       1813 rIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGl 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5962 CATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCA 6021

Qy       1833 nSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSe 1853
              ||||||||||:::                        |||   |||||||||||||||||
Db       6022 GAGCAGCGGGGCT------------------------GGGGGCAGCAGCCGCCCCGCCTC 6057
```

```
Qy       1853 rHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGl 1873
              ||||:::      |||||||||||||||||||||||||||||||||||||||||||||||
Db       6058 CCACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCA 6111

Qy       1873 nArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSe 1893
              |||||||||||||||||||||||||   |||||||||::::::|||:::|||||||||
Db       6112 GAGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGG 6171

Qy       1893 rLysProThrValLeu---------ArgSerThrSerThrSerSerProValArgProAl 1910
                  ||||||||||||         |||||||||||||||||||||||||||||||||||
Db       6172 CACGCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGC 6231

Qy       1910 aAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPr 1930
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       6232 TGCCACATTCCCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCC 6291

Qy       1930 oThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluAr 1950
              ||||||||||||||||||||||||||||||||||      ||||||||||||||||||||
Db       6292 TACCCTCATGGAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCG 6351

Qy       1950 gProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGl 1970
              |   |||   |||   |||||||||||||||   |||||||||   |||         ||
Db       6352 GGCCCGGGTGGACGCTGGCCATGCCTTTCTTACCAAACCCCCGGGCCGG---------GA 6402

Qy       1970 uProAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGl 1990
              ||||||||||||||||||||||   ||||||||||||   |||   ||||||   |||
Db       6403 GCCCGCCTCCTCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAG 6462

Qy       1990 yHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAs 2010
               |||      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6463 CCACACAGCCATCGCCCGCACCCCAGCAAAGAACCTTGCACCCCACCATGCCAGTCCGGA 6522

Qy       2010 pProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysPr 2030
              ||||||||||   |||   ||||||||||||   ||||||||||||||||||||||||||
Db       6523 CCCGCCGGCG---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACC 6579

Qy       2030 oPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSe 2049
              |||||||||||||||||||||||||||||||||||||||||||   |||:::   |||||
Db       6580 CTTTTCCATCCAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAG 6639

Qy       2049 rProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGl 2069
              ||||:::||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db       6640 CCCCGATGGGGTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGG 6699

Qy       2069 yLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLy 2089
              ||||   |||   ||||||||||||:::|||||||||||||||||||||||||||   ||
Db       6700 GCTCTCCAAACCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAA 6759

Qy       2089 sGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLe 2109
              |||||||||||||:::||||||      ||||||||||||||||||||||||||||||||
Db       6760 GCAGCCAGGCCCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACT 6819

Qy       2109 uProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHi 2129
              |||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||
Db       6820 GCCCGAGAGCCAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCA 6879

Qy       2129 sGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrAr 2149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6880 CCAGAGGGTGGTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACGCG 6939

Qy       2149 gHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAl 2169
              ||||||||||||||||||||||   |||||||||||||||||||||||||||||||||||
Db       6940 CCACCACCCGCAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGC 6999

Qy       2169 aSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAs 2189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7000 CAGCTGCCCTGTCCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGA 7059

Qy       2189 pHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluPr 2209
              |||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7060 CCATGGCACCCCAGCCCGGGGATCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACC 7119
```

```
Qy       2209 oAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGl 2229
              |:::||||||||||||||||||      ||||||   |||||||||||||||||||||||
Db       7120 CAGCAAAACATCGGTCCTGGGCAGCAGCGAGGATGCCATTGAGCCTGTGTCCCCACCAGA 7179

Qy       2229 uGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGl 2249
              ||||||||||||||||||||||:::||||||      |||||||||||||||||||||||
Db       7180 GGGCATGACTGAGCCAGGACATGCTCGGAGCACTGCGTACCCACTGCTGTATCGAGACGG 7239

Qy       2249 yGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProPr 2269
              |||||||   ||||||   |||||||||   :::||||||||||||||||||||||||||
Db       7240 GGAACAGGGCGAGCCC---AGGATGGGTCTAGAGTCTCCAGGCAACACCAGCCAGCCGCC 7296

Qy       2269 oAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGl 2289
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7297 AACCTTCTTCAGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGA 7356

Qy       2289 uIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnPr 2309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||||
Db       7357 GATCAACAAGAAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCC 7416

Qy       2309 oGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSe 2329
              |||||||||||||||||||||||||||||||||||||   ||||||||||||   |||||
Db       7417 TGGGACGGAAATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAG 7476

Qy       2329 rGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAl 2349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7477 CCAGGCGGTGCAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGC 7536

Qy       2349 aLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAs 2369
              |||||||||||||||||||||||||||||||   |||||||||   ||||||||||||||
Db       7537 ACTCATGGGTAAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAA 7596

Qy       2369 nProLeuAsnAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGl 2388
              ||||||||||||||||||||||||||||   ||||||||||||||||||   ||||||||
Db       7597 CCCTCTGAATGCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGG 7656

Qy       2388 yArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyAr 2408
              |||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db       7657 ACGGAGTGACCACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTCTGGCAG 7716

Qy       2408 gProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProPr 2428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7717 ACCTAGCAGCCGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCC 7776

Qy       2428 oSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnAr 2448
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7777 TTCTGTCTCCTCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCG 7836

Qy       2448 gValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIl 2468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7837 TGTGTGGGAGGACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGAT 7896

Qy       2468 eMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySe 2488
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db       7897 TATGAGGCTACAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAG 7956

Qy       2488 rGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSe 2508
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7957 CGGGCCCCTAGCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTC 8016

Qy       2508 rGlnTyrGluThrLeuSerAspSerGlu 2517
              ||||||||||||||||||||||||||||
Db       8017 ACAGTATGAGACACTCTCGGACAGCGAG 8044

RESULT 5
AF125671
LOCUS       AF125671                 8544 bp    mRNA    linear   ROD 04-APR-1999
DEFINITION  Mus musculus silencing mediator of retinoic acid and thyroid
            hormone receptor extended isoform (Smrte) mRNA, complete cds.
ACCESSION   AF125671
VERSION     AF125671.1  GI:4559295
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
```

```
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1  (bases 1 to 8544)
  AUTHORS Park,E.J., Schroen,D.J., Yang,M., Li,H., Li,L. and Chen,J.D.
  TITLE   SMRTe, a silencing mediator for retinoid and thyroid hormone
          receptors-extended isoform that is more related to the nuclear
          receptor corepressor
  JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)
  MEDLINE 99199215
   PUBMED 10097068
REFERENCE 2  (bases 1 to 8544)
  AUTHORS Chen,J.D.
  TITLE   Direct Submission
  JOURNAL Submitted (03-FEB-1999) Pharmacology and Molecular Toxicology,
          University of Massachusetts Medical School, 55 Lake Avenue North,
          Worcester, MA 01655, USA
FEATURES           Location/Qualifiers
     source        1. .8544
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /db_xref="taxon:10090"
     gene          1. .8544
                   /gene="Smrte"
     CDS           160. .7548
                   /gene="Smrte"
                   /function="transcriptional corepressor"
                   /note="SMRTe"
                   /codon_start=1
                   /product="silencing mediator of retinoic acid and thyroid
                   hormone receptor extended isoform"
                   /protein_id="AAD22972.1"
                   /db_xref="GI:4559296"
                   /translation="MSGSTQPVAQTWRAAEPRYPPHGISYPVQIARSHTDVGLLEYQH
                   HPRDYTSHLSPGSIIQPQRRRPSLLSEFQPGSERSQELHLRPESRTFLPELGKPDIEF
                   TESKRPRLELLPDTLLRPSPLLATGQPSGSEDLTKDRSLAGKLEPVSPPSPPHADPEL
                   ELAPSRLSKEELIQNRLDRVDREITMVEQQISKLKKKQQLEEEAAKPPEPEKPVSPP
                   PIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTRQYHENIKINQA
                   MRKKLILYFKRRNHARKQWEQRFCQRYDQLMEAWEKKVERIENNPRRRAKESKVREYY
                   EKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ
                   LAVIRHVVRRDQQRIKFINMNGLMDDPMKVYKDRQVTNMWSEQERDTFREKFMQHPKN
                   FGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQM
                   ARSSQEEKEEKEKEKEADKEEEKQDAENEKEELSKEKTDDTSGEDNDEKEAVASKGRK
                   TANSQGRRKGRITRSMANEANHEETATPQQSSELASMEMNESSRWTEEEMETAKKGLL
                   EHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKT
                   PAAASEETAFPPAAEDEEMEASGASANEEELAEEAEASQASGNEVPRVGECSGPAAVN
                   NSSDTESVPSPRSEATKDTGPKPTGTEALPAATQPPVPPPEEPAAAPAEPSPVPDASG
                   PPSPEPSPSPAAPPATVDKDEQEAPAAPAPQTEDAKEQKSEAEEIDVGKPEEPEASEE
                   PPESVKSDHKEETEEEPEDKAKGTEAIETVSEAPLKVEEAGSKAAVTKGSSSGATQDS
                   DSSATCSADEVDEPEGGDKGRLLSPRPSLLTPAGDPRASTSPQKPLDLKQLKQRAAAI
                   PPIVTKVHEPPREDTVPPKPVPPVPPPTQHLQPEGDVSQQSGGSPRGKSRSPVPPAEK
                   EAEKPAFFPAFPTEGQSYRLSPHAGHRLPSHPPREVIKTSTRADPLFSYTPPGHPLPL
                   GLHDSARPVLPRPPISNPPPLISSAKHPGVLERQLGAISQGMSVQLRVPHSEHAKPMG
                   PLTMELPLAVDPKKLGTALAPPPVEASPRASQYPGCRRPQLQRLYHPRTPADVLYKGT
                   ISRIVGEDSPSRLDRAREDTLPKGHVIYEGKKGHVLSYEGGMSVSQCSKEDGRSSSGP
                   PHETAAPKRTYDMMEGRVGRTVTSASIEGLMGRAIPEQHSPHLKEQHHIRGSITQGIP
                   RSYVEAQEDYLRREAKLLKREGTPPPPPPPPRDLTETYKPRPLDPLGPLKLKPTHEGVV
                   ATVKEAGRSIHEIPREELRRTPELPLAPRPLKEGSITQGTPLKYDSGAPSTGTKKHDV
                   RSIIGSPGRPFPALHPLDIMADARALERACYEESLKSRSGTSSGAGGSITRGAPVVVP
                   ELGKPRQSPLTYEDHGAPFTSHLPRGSPVTTREPTPRLQEGSLLSSKASQDRKLTSTP
                   REIAKSPHSTVPEHHPHPISPYEHLLRGVTGVDLYRGHIPLAFDPTSIPRGIPLEAAA
                   AAYYLPRHLAPSPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNAASAMA
                   QRADMLRGLSPRESSLALNYSAGPRGIIDLSQVPHLPVLVPPTPGTPATAIDRLAYLP
                   TAPPPFSSRHSSSPLSPGGPTHLAKPTATSSSERERERERERDKSILTSTTTVEHAPI
                   WRPGTEQSSGAGGSSRPASHTHQHSPISPRTQDALQQRPSVLHNTSMKGVVTSVEPGT
                   PTVLRSTSTSSPVRPAATFPPATHCPLGGTLEGVYPTLMEPVLLPKETSRVARPERPR
                   VDGGHAFLTKPPAREPASSPSKSSEPRSLAPPSSSHTAIARTPAKSLAPHHASPDPPG
                   PTSASDLHREKTQSKPFSIQELELRSLGYHSGAGYSPDGVEPISPVSSPSLTHDKGLS
                   KPLEELEKSHLEGELRHKQPGPMKLSAEAAHLPHLRPLPESQPSSSPLLQTAPGIKGH
                   QRVVTLAQHISEVITQDYTRHHPQQLSGPLPAPLYSFPGASCPVLDLRRPPSDLYLPP
                   PDHGTPARGSPHSEGGKRSPEPSKTSVLGSSEDAIEPVSPPEGMTEPGHARSAVYPLL
                   YRDGEQGEPRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEY
                   NIGQPGTEIFNMPAITGAGLMTCRSQAVQEHASTNMGLEAIIRKALMGKYDQWEEPPP
                   LGANAFNPLNASASLPAAAMPITTADGRSDHALTSPGGGGKAKVSGRPSSRKAKSPAP
                   GLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMA
                   SPPPPGLAAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE"
ORIGIN

Alignment Scores:
Pred. No.:              1.5e-210       Length:        8544
Score:                  10832.50        Matches:       2149
Percent Similarity:     87.94%         Conservative:  90
Best Local Similarity:  84.41%         Mismatches:    197
Query Match:            81.97%         Indels:        117
DB:                     10             Gaps:          33

US-09-522-753-5 (1-2517) x AF125671 (1-8544)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db        160 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 219

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db        220 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 279

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db        280 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 339

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        340 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGGTCTCAG 399

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||   ::::::||||||||||||||||||   ::::::
Db        400 GAGCTCCACCTGCGCCCTGAGTCCCGCACGTTCCTGCCTGAGCTGGGCAAGCCCGACATA 459

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||   ||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        460 GAATTCACCGAGAGCAAGCGCCCCCGCCTGGAGCTACTACCCGATACCCTGCTGCGCCCA 519

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        520 TCACCCCTGCTGGCCACTGGGCAGCCGAGTGGGTCTGAAGACCTTACCAAGGACCGTAGC 579

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        580 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 639

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsn---MetAspArgVal 179
              ||||||   |||   ||||||||||||||||||||||||||||||   :::|||||||||
Db        640 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACAGATTGGACCGCGTG 699

Qy        180 AspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGln 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 GACCGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAG 759

Qy        200 LeuGluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIle 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TTGGAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATA 819

Qy        220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGlu 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 GAATCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGAAGAAAGCCGAA 879

Qy        240 AlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGln 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 GCCGCACACCGGATCCTAGAAGGCCTGGGGCCCCAGGTGGAGCTGCCTCTGTACAACCAG 939

Qy        260 ProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLys 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CCGTCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAG 999

Qy        280 LeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCys 299
              |||||||||||||||||||||||||||||||||||||||||||||:::|||:::||||||
Db   1000 CTGATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGGAACAGCGCTTCTGC 1059

Qy    300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsn 319
          |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db   1060 CAGCGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAAT 1119

Qy    320 ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1120 CCGCGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAG 1179

Qy    340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySer 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1180 ATCCGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGT 1239

Qy    360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1240 GGGCTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTG 1299

Qy    380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
          ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db   1300 TCTGAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCC--GCCATGTTG 1357

Qy    400 TyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspPro 419
          ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db   1358 TACGACGC-GACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCC 1416

Qy    420 MetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThr 439
          |||||||||||||||||||||||||||   |||||||||||||||||||||::::::|||
Db   1417 ATGAAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACC 1476

Qy    440 PheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGlu 459
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 TTCCGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAG 1536

Qy    460 ArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyr 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 AGAAAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTAC 1596

Qy    480 LysSerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGln 499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 AAGAGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAG 1656

Qy    500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
          ||||||||||||||||||||||||            |||   ||||||||||||||||||
Db   1657 CAACAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAG 1704

Qy    520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
          |||:::||||||||||||||||||||||||:::|||||||||||||||   :::   |||
Db   1705 AAGGAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAG 1764

Qy    540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
          |||:::||||||:::|||   |||||||||||||||||||||||||||||||||||||||
Db   1765 AACGAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACGAT 1824

Qy    560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1825 GAGAAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAA 1884

Qy    580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
          ||||||||||||||||||||||||||||||||||||   ||||||      |||||||||
Db   1885 GGCCGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAG 1944

Qy    600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
          ||||||:::||||||||||||||||||:::||||||||||||||||||||||||||||||
Db   1945 CAAAGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAG 2004

Qy    620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2005 ATGGAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGC 2064

Qy    640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2065 ATGGTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2124

Qy    660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2125 CAGAACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCT 2184

Qy    680 ArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProVal 699
          ||||||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db   2185 CGGAGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCC 2244

Qy    700 ValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGlu 719
             |||||||||||||||||||||||||||   |||   ||||||||||||:::   |||
Db   2245 GCTGAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAG 2304

Qy    720 GluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGly 738
          ||||||||||||      ||||||||||||||||||||||||   |||||||||||||||
Db   2305 GAGGCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGC 2364

Qy    739 ProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAla 758
          ||||||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db   2365 CCAGCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCC 2424

Qy    759 AlaLysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAsp 776
             |||||||||         ||||||||||||      |||      |||   |||
Db   2425 ACGAAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCC 2475

Qy    777 GlyProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluPro 796
                ||||||   |||      ||||||         :::   ||||||   ||||||
Db   2476 ACCCAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGCAGCCCCTGCTGAGCCC 2529

Qy    797 ThrProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSer 816
          :::|||      :::|||:::|||   |||:::|||   |||:::   |||||||||:::
Db   2530 TCCCCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTCC---CCATCACCTGCC 2586

Qy    817 AlaProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProPro 836
          |||||||||      |||   |||:::|||:::|||         |||   ||||||
Db   2587 GCACCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAG 2646

Qy    837 ValGluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGly 856
             |||:::   :::|||||||||      |||   |||         :::|||   |||
Db   2647 ACAGAAGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGGA 2697

Qy    857 LysAlaGluGluPro---------------------------ValLysSerGluCysThr 867
          |||   |||||||||                           |||||||||:::
Db   2698 AAGCCAGAGGAGCCCGAGGCCTCTGAGGAGCCCCCGGAGAGTGTAAAGAGTGACCACAAG 2757

Qy    868 GluGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThr 886
          ||||||   ||||||   |||      |||   |||      ||||||   |||
Db   2758 GAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTG 2817

Qy    887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThr--- 905
          :::|||      ||||||   |||      |||   ||||||      |||   |||
Db   2818 TCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGTAGCAAGGCAGCTGTGACCAAG 2871

Qy    906 AlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGlu 925
                ||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db   2872 GGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACCTGCAGTGCCGATGAG 2931

Qy    926 ValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeu 945
          |||||||||   |||||||||||||||   ||||||||||||||||||||||||||||||
Db   2932 GTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTC 2991

Qy    946 ThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGln 965
          ||||||   |||||||||||||||:::   ||||||||||||||||||||||||||||||
Db   2992 ACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAG 3051

Qy    966 LeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProPro 985
          |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db   3052 CTGAAGCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAGGTCCATGAGCCCCCC 3108

Qy    986 ArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeu 1005
          |||||||||      |||   ||||||   ||||||   |||||||||   |||:::|||
Db   3109 CGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTA 3168
```

```
Qy   1006 GlnProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSer 1025
          |||||||||   |||      ||||||   |||   ||||||||||||||||||||||||
Db   3169 CAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGC 3228

Qy   1026 ProAlaProProAlaAspLysGlu------------------------AlaPheAlaAla 1037
          |||   |||||||||:::||||||                        ||||||
Db   3229 CCAGTGCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACT 3288

Qy   1038 GluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProVal 1057
          |||      |||||||||   :::||||||   |||:::|||         ||||||
Db   3289 GAGGGCCA-AAGCTACCGACTGAGCCCCCACGCTGGTCATCGGCT---GCCTTCCCAT-- 3342

Qy   1058 ProProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyr 1077
          |||||||||||||||||||||   |||      |||   ||||||      |||||||||
Db   3343 CCTCCACGGGAGGTGATCAAGACTTCCACACGCGCT---GACCCT---CTCTTCTCCTAC 3396

Qy   1078 AlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuPro 1097
             ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db   3397 ACACCCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCA 3456

Qy   1098 ArgProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerVal 1117
          |||||||||   ||||||||||||||||||||||||||||||||||||||||||   |||
Db   3457 CGTCCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTA 3513

Qy   1118 LeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyr 1137
          ||||||||||||:::|||||||||||||||||||||||||||||||||   ||||||:::
Db   3514 CTTGAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCAC 3573

Qy   1138 SerGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAsp 1157
          |||||||||||||||   |||:::||||||:::||||||   |||||||||   :::|||
Db   3574 TCAGAGCATGCCAAG---CCCATGGGCCCTCTCACCATGGAGCTGCCCCTTGCCGTGGAC 3630

Qy   1158 ProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGln 1177
          ||||||||||||
Db   3631 CCTAAGAAGCTG------------------------------------------------ 3642

Qy   1178 AlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGly 1197
                                                                   |||
Db   3643 ---------------------------------------------------------GGG 3645

Qy   1198 ThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVal 1217
          ||||||   ||||||         ||||||||||||||||||   ||||||||||||
Db   3646 ACAGCACT-GGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCT-CCCAGTACCCGGGCT 3703

Qy   1218 ProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVal 1237
                            :::|||||||||||||||||||||   |||||||||||||||
Db   3704 GCAGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGC-ACGCCCGCAGACGTC 3756

Qy   1238 LeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArg 1257
          ||||||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db   3757 CTCTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGG 3816

Qy   1258 GlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisVal 1277
             |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db   3817 GCACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTC 3876

Qy   1278 LeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1297
          |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db   3877 CTATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAGC 3936

Qy   1298 SerGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArg 1317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3937 TCGGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCGT 3996

Qy   1318 ValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPro 1337
          |||||||||   ::::::||||||||||||||||||||||||||||||||||||   |||
Db   3997 GTAGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CCT 4053

Qy   1338 GluArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGln 1357
          |||:::|||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db   4054 GAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAA 4110

Qy   1358 GlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4111 GGCATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGGAGGCCAAGCTC 4170

Qy   1378 LeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyr 1397
          ||||||||||||||||||||||||||||||||||||   |||||||||||||||   |||
Db   4171 TTGAAGCGAGAAGGGACACCACCTCCCCCACCACCACCTCGGGACCTGACTGAGACCTAC 4230

Qy   1398 LysThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeu 1414
          |||   :::            ||||||||||||||||||||||||   |||||||||:::
Db   4231 AAGCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGTG 4290

Qy   1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4291 GTAGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCGC 4350

Qy   1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4351 CGCACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGGC 4410

Qy   1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
          ||||||||||||||||||:::||||||   :::||||||:::||||||||||||||||||
Db   4411 ACCCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCGC 4470

Qy   1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
          |||:::||||||||||||||||||   ||||||   :::||||||||||||:::||||||
Db   4471 TCCATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCT 4530

Qy   1495 AspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThr 1514
          |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db   4531 GACGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACC 4590

Qy   1515 AlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGly 1534
          :::|||   :::||||||||||||   |||||||||||||||:::|||||||||||||||
Db   4591 AGCAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGC 4650

Qy   1535 LysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeu 1554
          ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db   4651 AAGCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTG 4710

Qy   1555 ProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeu 1574
          |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db   4711 CCACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCTC 4770

Qy   1575 SerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLys 1594
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4771 CTATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAAG 4830

Qy   1595 SerProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeu 1614
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4831 TCCCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTTG 4890

Qy   1615 LeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPro 1634
          ||||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||
Db   4891 CTCCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCCC 4950

Qy   1635 ThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArg 1653
          |||||||||||||||||||||||||||:::   |||||||||||||||||||||||||||
Db   4951 ACCTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCGG 5010

Qy   1654 HisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyr 1673
          ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db   5011 CACTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTAC 5070

Qy   1674 ProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGln 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5071 CCTGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAG 5130

Qy   1694 GlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeu 1713
          |||||||||||||||   |||:::||||||||||||||||||||||||||||||||||||
Db   5131 CAGATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTG 5190

Qy   1714 SerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAsp 1733
          |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db   5191 TCACCGCGAGAGTCCTCGCTGGCCCTCAATTATTCCGCTGGCCCAAGAGGCATTATCGAC 5250

Qy   1734 LeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThr 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5251 CTGTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACC 5310

Qy   1754 AlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSer 1773
          |||:::||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db   5311 GCCATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGT 5370

Qy   1774 SerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSer 1793
          ||||||||||||||||||||||||||||||||||||   |||||||||   |||||||||
Db   5371 AGCTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCA 5430

Qy   1794 SerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
          |||||||||||||||:::|||:::||||||||||||                  ||||||
Db   5431 TCGGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTCC 5472

Qy   1814 IleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGln 1833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5473 ATCCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAG 5532

Qy   1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSer 1853
          |||||||||:::                        |||   ||||||||||||||||||
Db   5533 AGCAGCGGGGCT------------------------GGGGGCAGCAGCCGCCCCGCCTCC 5568

Qy   1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGln 1873
          |||:::      ||||||||||||||||||||||||||||||||||||||||||||||||
Db   5569 CACACC------CACCAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCAG 5622

Qy   1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
          ||||||||||||||||||||||||   |||||||||::::::|||:::|||||||||   
Db   5623 AGGCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGC 5682

Qy   1894 LysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPhe 1913
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5683 ACGCCCACGGTCCTGAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGCCACATTC 5742

Qy   1914 ProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMet 1933
          ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db   5743 CCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTACCCTCATG 5802

Qy   1934 GluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAla 1953
          ||||||||||||||||||||||||      ||||||||||||||||||||||||||||||   
Db   5803 GAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGCCCCGTGTG 5862

Qy   1954 AspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSer 1973
          |||   |||||||||||||||   |||||||||||||||         ||||||||||||
Db   5863 GACGGTGGCCATGCCTTCCTCACCAAACCCCCGGCCCGG---------GAGCCCGCCTCC 5913

Qy   1974 SerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThr 1993
          ||||||||||||   ||||||||||||   |||   ||||||   |||   |||      
Db   5914 TCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCACACAGCC 5973

Qy   1994 IleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAla 2013
          |||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||   
Db   5974 ATCGCCCGCACCCCAGCAAAGAGCCTTGCACCCCACCATGCCAGTCCGGACCCGCCGGGG 6033

Qy   2014 ProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIle 2033
             |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db   6034 ---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTTTTCCATC 6090

Qy   2034 GlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSerProGluGly 2052
          |||||||||||||||||||||||||||||||||   |||:::   |||||||||:::|||
Db   6091 CAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAGCCCCGATGGG 6150

Qy   2053 ValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLys 2072
          |||||||||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db   6151 GTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGGGCTCTCCAAA 6210

Qy   2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
             ||||||||||||:::|||||||||||||||||||||||||||   ||||||||||||
Db   6211 CCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAAGCAGCCAGGC 6270

Qy   2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSer 2112
          |||:::||||||      ||||||||||||||||||||||||||||||||||||||||||
Db   6271 CCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACTGCCCGAGAGC 6330

Qy   2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVal 2132
          |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db   6331 CAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCAGAGGGTG 6390

Qy   2133 ValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisPro 2152
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6391 GTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACCCGGCACCACCCG 6450

Qy   2153 GlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysPro 2172
          ||||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db   6451 CAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGCCAGCTGCCCT 6510

Qy   2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGlyAla 2192
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
Db   6511 GTGCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGACCATGGCACC 6570

Qy   2193 ProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThr 2212
          |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db   6571 CCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACCCAGCAAAACA 6630

Qy   2213 SerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMetThr 2232
          ||||||||||||      |||||||||   |||||||||||||||||||||||||||||||||
Db   6631 TCGGTCCTGGGCAGCAGTGAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGGCATGACT 6690

Qy   2233 GluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThr 2252
          ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||   
Db   6691 GAGCCAGGACATGCTCGGAGCGCTGTGTACCCACTGCTGTATCGAGACGGGGAACAGGGC 6750

Qy   2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePhe 2272
          ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6751 GAGCCC---AGGATGGGCTCTAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTC 6807

Qy   2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLys 2292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6808 AGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGATCAACAAG 6867

Qy   2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGlu 2312
          ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db   6868 AAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCCTGGGACGGAA 6927

Qy   2313 IlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332
          |||||||||||||||||||||||||||   ||||||||||||   |||||||||||||||
Db   6928 ATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAGCCAGGCGGTG 6987

Qy   2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6988 CAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGCACTCATGGGT 7047

Qy   2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
          |||||||||||||||||||||   |||||||||   ||||||||||||||||||||||||
Db   7048 AAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAACCCTCTGAAT 7107

Qy   2373 AlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
          ||||||||||||||||||   ||||||||||||||||||   ||||||||||||||||||
Db   7108 GCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGGACGGAGTGAC 7167

Qy   2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7168 CACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTCTGGCAGACCTAGCAGC 7227

Qy   2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7228 CGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCCTTCTGTCTCC 7287

Qy   2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7288 TCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCGTGTGTGGGAG 7347
```

```
Qy       2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7348 GACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGATTATGAGGCTA 7407

Qy       2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       7408 CAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAGCGGGCCCCTA 7467

Qy       2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7468 GCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTCACAGTATGAG 7527

Qy       2512 ThrLeuSerAspSerGlu 2517
              ||||||||||||||||||
Db       7528 ACACTCTCGGACAGCGAG 7545

RESULT 6
AF113002
LOCUS       AF113002                7465 bp    mRNA    linear   ROD 20-MAR-1999
DEFINITION  Mus musculus silencing mediator of retinoic acid and thyroid
            hormone receptor beta mRNA, complete cds.
ACCESSION   AF113002
VERSION     AF113002.1  GI:4454549
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 7465)
  AUTHORS   Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and
            Evans,R.M.
  TITLE     Unique forms of human and mouse nuclear receptor corepressor SMRT
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)
  MEDLINE   99178941
   PUBMED   10077563
REFERENCE   2  (bases 1 to 7465)
  AUTHORS   Downes,M.R., Ordentlich,P. and Evans,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk
            Institute for Biological Studies, 10010 North Torrey Pines Road, La
            Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1..7465
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="spleen; brain"
     CDS             363..7124
                     /function="transcriptional co-repressor"
                     /note="mSMRT beta; longer SMRT alpha isoform encoded by
                     GenBank Accession Number AF113001"
                     /codon_start=1
                     /product="silencing mediator of retinoic acid and thyroid
                     hormone receptor beta"
                     /protein_id="AAD20945.1"
                     /db_xref="GI:4454550"
                     /translation="MSGSTQPVAQTWRAAEPRYPPHGISYPVQIARSHTPLYNQPSDT
                     RQYHENIKINQAMRKKLILYFKRRNHARKQWEQRFCQRYDQLMEAWEKKVERIENNPR
                     RRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGL
                     SEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMDDPMKVYKDRQVTNMWSEQER
                     DTFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQ
                     QQQQQQQQQQMARSSQEEKEEKEKEKEADKEEEKQDAENEKEELSKEKTDDTSGED
                     NDEKEAVASKGRKTANSQGRRKGRITRSMANEANHEETATPQQSSELASMEMNESSRW
                     TEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKM
                     EKERNARRKKKKTPAAASEETAFPPAAEDEEMEASGASANEEELAEEAEASQASGNEV
                     PRVGECSGPAAVNNSSDTESVPSPRSEATKDTGPKPTGTEALPAATQPPVPPPEEPAV
                     APAEPSPVPDASGPPSPEPSHHLPHPRLLWTRMNKKPRLLQLPRQRMPRSRSLRPRRS
                     MWEKPEEPEASEKPPKSVKSDHKKETEEEPEDKAKGTEAIETVSEAPLKVEKAGSKAA
                     VTKGSSSGATQDSDSSATCSADEVDEPEGGDKGRLLSPRPSLLTPAGDPRASTSPQKP
                     LDLKQLKQRAAAIPPIVTKVHEPPREDTVPPKPVPPVPPPTQHLQPEGDVSQQSGGSP
                     RGKSRSPVPPAEKEAEKPAFFPAFPTEGPKLPTEPPRWSSGLPFPIPPREVIKTSPHA
                     ADPSAFSYTPPGHPLPLGLHDSARPVLPRPPISNPPPLISSAKHPGVLERQLGAISQQ
                     GMSVQLRVPHSEHAKAPMGPLTMGLPLAVDPKKLGTALGSATSGSITKGLPSTRAADG
                     PSYRGSITHGTPADVLYKGTISRIVGEDSPSRLDRAREDTLPKGHVIYEGKKGHVLSY
                     EGGMSVSQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRTVTSASIEGLMGRAIPEQ
                     HSPHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPPRDLTETYK
                     PRPLDPLGPLKLKPTHEGVVATVKEAGRSIHEIPREELRRTPELPLAPRPLKEGSITQ
                     GTPLKYDSGAPSTGTKKHDVRSIIGSPGRPFPALHPLDIMADARALERACYEESLKSR
                     SGTSSGAGGSITRGAPVVVPELGKPRQSPLTYEDHGAPFTSHLPRGSPVTTREPTPRL
                     QEGSLLSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVTGVDLYRGH
                     IPLAFDPTSIPRGIPLEAAAAAYYLPRHLAPSPTYPHLYPPYLIRGYPDTAALENRQT
                     IINDYITSQQMHHNAASAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPV
                     LVPPTPGTPATAIDRLAYLPTAPPPFSSRHSSSPLSPGGPTHLAKPTATSSSERERER
                     ERERDKSILTSTTTVEHAPIWRPGTEQSSGAGGSSRPASHTHQHSPISPRTQDALQQR
                     PSVLHNTSMKGVVTSVEPGTPTVLRWARSTSTSSPVRPAATFPPATHCPLGGTLEGVY
                     PTLMEPVLLPKETSRVARPERARVDAGHAFLTKPPGREPASSPSKSSEPRSLAPPSSS
                     HTAIARTPAKNLAPHHASPDPPAPTSASDLHREKTQSKPFSIQELELRSLGYHSGAGY
                     SPDGVEPISPVSSPSLTHDKGLSKPLEELEKSHLEGELRHKQPGPMKLSAEAAHLPHL
                     RPLPESQPSSSPLLQTAPGIKGHQRVVTLAQHISEVITQDYTRHHPQQLSGPLPAPLY
                     SFPGASCPVLDLRRPPSDLYLPPPDHGTPARGSPHSEGGKRSPEPSKTSVLGSSEDAI
                     EPVSPPEGMTEPGHARSTAYPLLYRDGEQGEPRMGLESPGNTSQPPTFFSKLTESNSA
                     MVKSKKQEINKKLNTHNRNEPEYNIGQPGTEIFNMPAITGAGLMTCRSQAVQEHASTN
                     MGLEAIIRKALMGKYDQWEEPPPLGANAFNPLNASASLPAAAMPITTADGRSDHALTS
                     PGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRP
                     SSAGSTPFPYNPLIMRLQAGVMASPPPPGLAAGSGPLAGPHHAWDEEPKPLLCSQYET
                     LSDSE"
ORIGIN

Alignment Scores:
Pred. No.:              1.16e-191      Length:         7465
Score:                  9894.00        Matches:        1955
Percent Similarity:     79.99%         Conservative:   84
Best Local Similarity:  76.70%         Mismatches:     183
Query Match:            74.87%         Indels:         328
DB:                     10             Gaps:           30

US-09-522-753-5 (1-2517) x AF113002 (1-7465)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db        363 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 422

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||   :::||||||||||||||||||||||||:::||||||
Db        423 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACG--------------- 467

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60

Db        467 ------------------------------------------------------------ 467

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80

Db        467 ------------------------------------------------------------ 467

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100

Db        467 ------------------------------------------------------------ 467

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120

Db        467 ------------------------------------------------------------ 467

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140

Db        467 ------------------------------------------------------------ 467

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160

Db        467 ------------------------------------------------------------ 467

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180

Db        467 ------------------------------------------------------------ 467

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
```

```
Db           467 ------------------------------------------------------------ 467

Qy           201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220

Db           467 ------------------------------------------------------------ 467

Qy           221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240

Db           467 ------------------------------------------------------------ 467

Qy           241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                                                           ||||||||||||||||||
Db           468 ------------------------------------------CCTCTGTACAACCAGCCG 485

Qy           261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           486 TCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTG 545

Qy           281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                 ||||||||||||||||||||||||||||||||||||||||||:::|||:::|||||||||
Db           546 ATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGGAACAGCGCTTCTGCCAG 605

Qy           301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                 ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db           606 CGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAATCCG 665

Qy           321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           666 CGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAGATC 725

Qy           341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           726 CGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGTGGG 785

Qy           361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           786 CTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTGTCT 845

Qy           381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           846 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGTTGTAC 905

Qy           401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db           906 GACGCGGACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCCATG 965

Qy           421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                 ||||||||||||||||||||||||   |||||||||||||||||||||::::::||||||
Db           966 AAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACCTTC 1025

Qy           441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1026 CGTGAGAAGTTTATGCAGCACCCTAAGAACTTTGGCCTGATTGCCTCATTCCTGGAGAGA 1085

Qy           461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1086 AAGACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTACAAG 1145

Qy           481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1146 AGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGCAGCAGCAGCAGCAGCAA 1205

Qy           501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
                 |||||||||||||||||||||            |||   |||||||||||||||||||||
Db          1206 CAACAGCAGCAGCAGCAGCAG------------ATGGCACGGAGCAGCCAGGAGGAGAAG 1253

Qy           521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                 :::||||||||||||||||||||||||:::|||||||||||||||   :::   ||||||
Db          1254 GAGGAGAAGGAGAAGGAGAAGGAGGCCGACAAGGAGGAAGAGAAGCAGGATGCGGAGAAC 1313

Qy           541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                 :::||||||:::|||   ||||||||||||||||||||||||||||||||||||||||||
Db          1314 GAGAAGGAAGAACTCAGCAAGGAGAAGACAGACGACACTTCTGGCGAGGACAACGATGAG 1373
```

```
Qy           561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1374 AAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAAGGC 1433

Qy           581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                 |||||||||||||||||||||||||||||||||   ||||||      ||||||||||||
Db          1434 CGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAGCAA 1493

Qy           601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                 |||:::||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db          1494 AGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAGATG 1553

Qy           621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1554 GAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGCATG 1613

Qy           641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1614 GTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1673

Qy           661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1674 AACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCTCGG 1733

Qy           681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                 |||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db          1734 AGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCCGCT 1793

Qy           701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                 |||||||||||||||||||||||||||   |||   ||||||||||||:::   ||||||
Db          1794 GAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAGGAG 1853

Qy           721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArg---GlyGluCysSerGlyPro 739
                 |||||||||      ||||||||||||||||||||||||   ||||||||||||||||||
Db          1854 GCAGAAGCCTCACAGGCCTCTGGGAATGAGGTTCCCAGAGTTGGGGAGTGCAGTGGCCCA 1913

Qy           740 AlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAla 759
                 |||   |||||||||||||||||||||||||||:::|||||||||   :::||||||
Db          1914 GCTGCTGTCAACAACAGCTCTGATACTGAGAGTGTCCCATCCCCGCGTTCAGAAGCCACG 1973

Qy           760 LysAspThrGlyGlnAsnGlyProLysProProAlaThr------LeuGlyAlaAspGly 777
                 |||||||||         ||||||||||||      |||      |||   |||
Db          1974 AAGGACACT---------GGGCCTAAACCCACTGGCACTGAAGCATTGCCCGCTGCCACC 2024

Qy           778 ProProProGlyProProThrProProArgArgThrSerArgAlaProIleGluProThr 797
                    ||||||   |||      ||||||         :::   ||||||   ||||||:::
Db          2025 CAGCCACCTGTTCCT------CCTCCAGAAGAACCGGCAGTAGCCCCTGCTGAGCCCTCC 2078

Qy           798 ProAlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAla 817
                 |||      :::|||:::|||   |||:::|||   |||      |||||||||:::|||
Db          2079 CCAGTCCCTGATGCCAGTGGCCCACCATCCCCAGAGCCTTC----CCATCACCTGCCGCA 2134

Qy           818 ProProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProVal 837
                 ||||||      |||   |||:::|||:::|||         |||   ||||||
Db          2135 CCCCCGGCTACTGTGGACAAGGATGAACAAGAAGCCCCGGCTGCTCCAGCTCCCCAGACA 2194

Qy           838 GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLys 857
                 |||:::   :::|||||||||      |||   |||         :::|||   ||||||
Db          2195 GAGGATGCCAAGGAGCAGAAGTCTGAGGCCGAGGAG---------ATCGATGTGGAAAAA 2245

Qy           858 Ala-GluGluPro---------------------------ValLysSerGluCysThrGl 868
                 ||| |||||||||                           |||||||||:::      ::
Db          2246 GCCAGAGGAGCCCGAGGCTTCTGAGAAGCCCCCGAAGAGTGTAAAGAGTGACCACAAGAA 2305

Qy           868 uGluAlaGluGluGlyProAla---LysGlyLysAspAlaGluAlaAlaGluAlaThrAl 887
                 :|||   ||||||   |||      |||   |||      ||||||   |||      ::
Db          2306 GGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGCACAGAGGCCATTGAAACTGTGTC 2365

Qy           887 aGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLy 907
                 :|||      ||||||   |||   |||   |||               |||
Db          2366 TGAGGCACCACTTAAGGTGGAG---AAGGCTGGTAGCAAGGCAGCTGTGACCAAGGGTTC 2422
```

```
Qy         907 sSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluValAs 927
                ||||||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        2423 CAGCTCAGGTGCCACCCAGGACAGTGACTCCAGTGCCACCTGCAGTGCCGATGAGGTGGA 2482

Qy         927 pGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThrPr 947
               ||||   |||||||||||||||   |||||||||||||||||||||||||||||||||||
Db        2483 CGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTGTCACCAAGGCCCAGCCTCCTCACCCC 2542

Qy         947 oThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLy 967
               |   |||||||||||||||:::   |||||||||||||||||||||||||||||||||||
Db        2543 GGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAGAAGCCGCTGGACCTGAAGCAGCTGAA 2602

Qy         967 sGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArgGl 987
               ||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db        2603 GCAGCGAGCAGCCGCCATCCCCCCTATC---GTCACCAAGGTCCATGAGCCCCCCCGGGA 2659

Qy         987 uAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGlnPr 1007
               ||||      |||   ||||||   ||||||   |||||||||   ||||:::||||||||
Db        2660 GGACACAGTACCCCCAAAGCCAGTTCCCCCTGTGCCTCCACCCACGCAGCACCTACAGCC 2719

Qy        1007 oGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAl 1027
               ||||   |||      ||||||   |||   |||||||||||||||||||||||||||
Db        2720 AGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGTCCACGTGGCAAGTCCCGCAGCCCAGT 2779

Qy        1027 aProProAlaAspLysGlu------------------------AlaPheAlaAlaGluAl 1039
                |||||||||:::||||||                        ||||||      |||
Db        2780 GCCTCCTGCCGAGAAAGAGGCAGAGAAACCCGCATTCTTTCCGGCTTTCCCAACTGAGGG 2839

Qy        1039 aGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPr 1059
                   |||||||||   :::||||||   |||:::||||||||||||||||||:::|||||
Db        2840 CCCAAAGCTACCGACTGAGCCCCCACGCTGGTCATCGGGCCTGCCCTTCCCCATCCCTCC 2899

Qy        1059 oArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPr 1079
               ||||||||||||||||   ||||||||||||   |||||||||||||||||||||   ||
Db        2900 ACGGGAGGTGATCAAGACTTCCCCACACGCCGCTGACCCCTCTGCCTTCTCCTACACACC 2959

Qy        1079 oProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPr 1099
               ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db        2960 CCCCGGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCACGTCC 3019

Qy        1099 oProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGl 1119
               ||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db        3020 CCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTACTTGA 3076

Qy        1119 uArgGlnIleGlyAlaIleSer---GlnGlyMetSerValGlnLeuHisValProTyrSe 1138
               |||||||:::||||||||||||   |||||||||||||||||||||   ||||||:::||
Db        3077 GAGGCAGCTGGGTGCCATCTCCCAGCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCACTC 3136

Qy        1138 rGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPr 1158
               |||||||||||||||||||:::||||||:::||||||||||||||||||   :::|||||
Db        3137 AGAGCATGCCAAGGCCCCCATGGGCCCTCTCACCATGGGGCTGCCCCTTGCCGTGGACCC 3196

Qy        1158 oLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAl 1178
               ||||||||||
Db        3197 TAAGAAGCTG-------------------------------------------------- 3206

Qy        1178 aGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyTh 1198
                                                                      |||||
Db        3207 -------------------------------------------------------GGGAC 3211

Qy        1198 rAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPr 1218
               |||||||||||||         ||||||||||||||||||:::||||||||||||
Db        3212 AGCACTGGGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCTCCCCAGTACCCGGGCTGC 3271

Qy        1218 oSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLe 1238
                               :::|||||||||||||||||||||||||||||||||||||||||
Db        3272 AGACGGCCCC------AGCTACAGAGGCTCTATCACCCACGGCACGCCCGCAGACGTCCT 3325

Qy        1238 uTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGl 1258
               ||||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db        3326 CTACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGGGC 3385

Qy        1258 yArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLe 1278
                |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db        3386 ACGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTCCT 3445

Qy        1278 uSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSe 1298
               |||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db        3446 ATCCTATGAAGGTGGTATGTCCGTGTCACAGTGCTCTAAGGAGGATGGAAGGAGCAGCTC 3505

Qy        1298 rGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVa 1318
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3506 GGGCCCACCCCATGAGACTGCCGCCCCTAAACGCACCTATGACATGATGGAGGGCCGTGT 3565

Qy        1318 lGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGl 1338
               |||||||   ::::::||||||||||||||||||||||||||||||||||||   |||||
Db        3566 AGGCAGGACTGTCACCTCAGCCAGCATAGAGGGACTCATGGGCCGCGCCATC---CCTGA 3622

Qy        1338 uArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGl 1358
               |:::|||||||||   ||||||||||||||||||||||||||||||||||||||||||||
Db        3623 GCAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAAGG 3679

Qy        1358 yIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLe 1378
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3680 CATCCCGAGGTCCTATGTGGAGGCGCAGGAGGACTACTTACGGCGGGAGGCCAAGCTCTT 3739

Qy        1378 uLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLy 1398
               ||||||||||||||||||||||||||||||||||   |||||||||||||||   |||||
Db        3740 GAAGCGAGAAGGGACACCACCACCCCCACCACCACCTCGGGACCTGACTGAGACCTACAA 3799

Qy        1398 sThrGln---------AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuVa 1415
               |   :::            ||||||||||||||||||||||||   |||||||||:::||
Db        3800 GCCCCGGCCCCTGGACCCTCTGGGTCCCCTGAAGCTGAAGCCGACTCACGAGGGTGTGGT 3859

Qy        1415 lAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHi 1435
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3860 AGCAACTGTGAAGGAGGCGGGCCGCTCTATCCATGAGATCCCGAGAGAGGAGCTGCGCCG 3919

Qy        1435 sThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyTh 1455
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3920 CACACCTGAGCTACCCCTGGCACCACGGCCTCTGAAGGAGGGTTCCATCACCCAGGGCAC 3979

Qy        1455 rProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSe 1475
               ||||||||||||||||:::||||||   :::||||||:::||||||||||||||||||||
Db        3980 CCCACTCAAGTACGACTCTGGGGCACCCTCCACTGGCACCAAGAAACACGACGTGCGCTC 4039

Qy        1475 rLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAs 1495
               |:::||||||||||||||||||   ||||||   :::||||||||||||:::||||||||
Db        4040 CATCATCGGCAGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCTGA 4099

Qy        1495 pAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAl 1515
               |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||::
Db        4100 CGCCCGGGCACTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACCAG 4159

Qy        1515 aSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLy 1535
               :|||   :::||||||||||||   |||||||||||||||:::|||||||||||||||||
Db        4160 CAGTGGTGCAGGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGCAA 4219

Qy        1535 sProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuPr 1555
               ||||||||||||||||||||||||||||||||||||||||||||||      ||||||||
Db        4220 GCCACGGCAAAGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTGCC 4279

Qy        1555 oArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSe 1575
               |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        4280 ACGTGGCTCCCCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGGCAGCCTCCT 4339

Qy        1575 rSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSe 1595
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4340 ATCCAGCAAGGCGTCCCAGGACCGGAAGCTGACATCTACACCCCGGGAGATCGCCAAGTC 4399

Qy        1595 rProHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLe 1615
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4400 CCCACACAGCACTGTGCCCGAGCACCACCCTCACCCCATCTCCCCCTATGAGCACTTGCT 4459

Qy        1615 uArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProTh 1635
               ||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||||
Db       4460 CCGGGGCGTGACTGGTGTGGACCTGTACCGTGGTCACATCCCATTGGCCTTTGACCCCAC 4519

Qy       1635 rSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHi 1654
              |||||||||||||||||||||||||:::   |||||||||||||||||||||||||||||
Db       4520 CTCCATACCCCGAGGGATCCCTCTGGAAGCAGCAGCCGCAGCCTACTACCTGCCCCGGCA 4579

Qy       1654 sLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrPr 1674
              ||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db       4580 CTTGGCCCCCAGCCCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTACCC 4639

Qy       1674 oAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGl 1694
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4640 TGACACGGCGGCCCTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAGCA 4699

Qy       1694 nMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSe 1714
              |||||||||||||   |||:::||||||||||||||||||||||||||||||||||||||
Db       4700 GATGCACCACAACGCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTGTC 4759

Qy       1714 rProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLe 1734
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4760 ACCGCGAGAGTCCTCGCTGGCCCTCAATTATGCCGCTGGCCCAAGAGGCATTATCGACCT 4819

Qy       1734 uSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAl 1754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4820 GTCCCAAGTGCCACACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACCGC 4879

Qy       1754 aMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSe 1774
              |:::||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db       4880 CATCGACCGCCTTGCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGTAG 4939

Qy       1774 rSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSe 1794
              ||||||||||||||||||||||||||||||||||   |||||||||   |||||||||||
Db       4940 CTCACCGCTGTCCCCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCATC 4999

Qy       1794 rGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIl 1814
              |||||||||||||:::|||:::||||||||||||                  ||||||||
Db       5000 GGAGCGGGAACGGGAACGTGAGCGGGAACGAGAC------------------AAGTCCAT 5041

Qy       1814 eLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSe 1834
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5042 CCTCACGTCTACCACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAGAG 5101

Qy       1834 rSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHi 1854
              |||||||:::                        |||   ||||||||||||||||||||
Db       5102 CAGCGGGCT-------------------------GGGGGCAGCAGCCGCCCCGCCTCCCA 5137

Qy       1854 sSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnAr 1874
              |:::      ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5138 CACC------CACCAGCACTCGCCCATCTCCCCCGGACCCAGGACGCCTTGCAGCAGAG 5191

Qy       1874 gProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLy 1894
              ||||||||||||||||||||||   |||||||||::::::|||:::|||||||||     
Db       5192 GCCCAGTGTGCTGCACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGCAC 5251

Qy       1894 sProThrValLeu---------ArgSerThrSerThrSerSerProValArgProAlaAl 1911
               ||||||||||||         ||||||||||||||||||||||||||||||||||||||
Db       5252 GCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGC 5311

Qy       1911 aThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProTh 1931
              |||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db       5312 CACATTCCCACCTGCCACCCACTGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTAC 5371

Qy       1931 rLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgPr 1951
              |||||||||||||||||||||||||||||||      |||||||||||||||||||||  
Db       5372 CCTCATGGAGCCCGTCCTGTTACCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGGC 5431

Qy       1951 oArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluPr 1971
               |||   |||   |||||||||||||||   |||||||||   |||         |||||
Db       5432 CCGGGTGGACGCTGGCCATGCCTTTCTTACCAAACCCCCGGGCCGG---------GAGCC 5482

Qy       1971 oAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHi 1991
              |||||||||||||||||||   ||||||||||||   |||   ||||||   |||   ||
Db       5483 CGCCTCCTCACCCAGCAAGAGCTCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCA 5542

Qy       1991 sAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspPr 2011
              |      |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5543 CACAGCCATCGCCCGCACCCCAGCAAAGAACCTTGCACCCCACCATGCCAGTCCGGACCC 5602

Qy       2011 oProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPh 2031
              |||||||   |||   ||||||||||||   |||||||||||||||||||||||||||||
Db       5603 GCCGGCG---CCCACCTCGGCCTCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTT 5659

Qy       2031 eSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHis---GlySerSerTyrSerPr 2050
              ||||||||||||||||||||||||||||||||||||||||   |||:::   ||||||||
Db       5660 TTCCATCCAGGAATTGGAACTCCGTTCTCTGGGTTACCACAGTGGAGCTGGCTACAGCCC 5719

Qy       2050 oGluGlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLe 2070
              |:::||||||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db       5720 CGATGGGGTGGAGCCCATCAGCCCGGTGAGCTCCCCCAGCCTGACCCACGACAAGGGGCT 5779

Qy       2070 uProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGl 2090
              |   |||   ||||||||||||:::|||||||||||||||||||||||||||   |||||
Db       5780 CTCCAAACCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAGGGGAGCTGCGGCACAAGCA 5839

Qy       2090 nProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuPr 2110
              ||||||||||:::||||||      |||||||||||||||||||||||||||||||||||
Db       5840 GCCAGGCCCCATGAAGCTCAGCGCGGAGGCTGCCCATCTCCCACATCTGCGGCCACTGCC 5899

Qy       2110 oGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGl 2130
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       5900 CGAGAGCCAGCCCTCATCCAGCCCACTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCA 5959

Qy       2130 nArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHi 2150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5960 GAGGGTGGTCACCCTGGCTCAGCACATCAGCGAGGTCATTACGCAGGACTACACGCGCCA 6019

Qy       2150 sHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSe 2170
              |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db       6020 CCACCCGCAGCAGCTCAGTGGCCCCCTTCCCGCCCCTCTCTACTCCTTTCCCGGAGCCAG 6079

Qy       2170 rCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHi 2190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6080 CTGCCCTGTCCTGGATCTTCGCCGCCCACCCAGTGACCTCTACCTCCCACCCCCCGACCA 6139

Qy       2190 sGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAs 2210
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||::
Db       6140 TGGCACCCCAGCCCGGGGATCCCCCCACAGTGAAGGGGGCAAAAGGTCCCCAGAACCCAG 6199

Qy       2210 nLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGl 2230
              :||||||||||||||||||      ||||||   ||||||||||||||||||||||||||
Db       6200 CAAAACATCGGTCCTGGGCAGCAGCGAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGG 6259

Qy       2230 yMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGl 2250
              |||||||||||||||||||:::||||||      ||||||||||||||||||||||||||
Db       6260 CATGACTGAGCCAGGACATGCTCGGAGCACTGCGTACCCACTGCTGTATCGAGACGGGGA 6319

Qy       2250 uGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAl 2270
              ||||   ||||||   |||||||||   :::|||||||||||||||||||||||||||  
Db       6320 ACAGGGCGAGCCC---AGGATGGGTCTAGAGTCTCCAGGCAACACCAGCCAGCCGCCAAC 6376

Qy       2270 aPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIl 2290
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6377 CTTCTTCAGTAAGCTGACTGAGAGCAACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGAT 6436

Qy       2290 eAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGl 2310
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
Db       6437 CAACAAGAAACTCAACACCCACAACCGGAACGAGCCAGAATACAATATTGGCCAGCCTGG 6496

Qy       2310 yThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGl 2330
              ||||||||||||||||||||||||||||||||||   ||||||||||||   ||||||||
Db       6497 GACGGAAATCTTCAACATGCCCGCCATCACTGGAGCAGGCCTTATGACCTGTAGAAGCCA 6556

Qy       2330 nAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLe 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6557 GGCGGTGCAAGAACACGCCAGCACCAACATGGGGCTAGAGGCCATTATTAGAAAGGCACT 6616
```

```
Qy         2350 uMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnPr 2370
                ||||||||||||||||||||||||||||   |||||||||   |||||||||||||||||
Db         6617 CATGGGTAAATATGATCAGTGGGAAGAGCCCCCGCCGCTCGGCGCCAATGCTTTTAACCC 6676

Qy         2370 oLeuAsnAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyAr 2389
                |||||||||||||||||||||||||   ||||||||||||||||||   |||||||||||
Db         6677 TCTGAATGCCAGCGCCAGTCTGCCCGCTGCTGCTATGCCCATAACCACTGCTGACGGACG 6736

Qy         2389 gSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgPr 2409
                ||||||||||   |||||||||||||||||||||||||||||||||||||||||||||||
Db         6737 GAGTGACCACGCACTCACCTCGCCAGGTGGAGGTGGGAAAGCCAAGGTCTCTGGCAGACC 6796

Qy         2409 oSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSe 2429
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6797 TAGCAGCCGAAAAGCCAAGTCGCCAGCACCAGGCCTAGCGTCCGGAGACCGACCCCCTTC 6856

Qy         2429 rValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgVa 2449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6857 TGTCTCCTCAGTACACTCAGAGGGGGACTGCAATCGCCGAACACCACTCACCAACCGTGT 6916

Qy         2449 lTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMe 2469
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         6917 GTGGGAGGACCGGCCCTCATCTGCAGGGTCCACGCCATTCCCCTACAACCCTTTGATTAT 6976

Qy         2469 tArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGl 2489
                ||||||||||||||||||||||||||||||||||||||||||||||   |||||||||||
Db         6977 GAGGCTACAGGCAGGTGTCATGGCCTCCCCGCCCCCACCTGGCCTTGCGGCAGGCAGCGG 7036

Qy         2489 yProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGl 2509
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         7037 GCCCCTAGCTGGTCCCCACCACGCCTGGGATGAGGAGCCCAAGCCACTGCTGTGTTCACA 7096

Qy         2509 nTyrGluThrLeuSerAspSerGlu 2517
                |||||||||||||||||||||||||
Db         7097 GTATGAGACACTCTCGGACAGCGAG 7121

RESULT 7
CQ722208
LOCUS       CQ722208                6339 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 8142 from Patent WO02068579.
ACCESSION   CQ722208
VERSION     CQ722208.1  GI:42283065
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 8142 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .6339
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.39e-189       Length:         6339
Score:                  9776.50         Matches:        1958
Percent Similarity:     75.36%          Conservative:   6
Best Local Similarity:  75.13%          Mismatches:     59
Query Match:            73.98%          Indels:         583
DB:                     6               Gaps:           15

US-09-522-753-5 (1-2517) x CQ722208 (1-6339)

Qy            1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db            1 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy           21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy           41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 180

Qy           61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 240

Qy           81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 300

Qy          101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 360

Qy          121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 420

Qy          141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy          161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy          181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy          201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 720

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 780

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 900

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db          901 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          961 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1020

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAG---AGGGTGGGCCAGCGGGGCAGTGGG 1077

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1078 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1137
```

```
Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1138 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1197

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1198 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGCTTATGGCCGACCCCATG 1257

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1258 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1317

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1318 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1377

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              |||   |||:::   |||
Db       1378 AAG---GTGAGTCGCTGC------------------------------------------ 1392

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500

Db       1392 ------------------------------------------------------------ 1392

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520

Db       1392 ------------------------------------------------------------ 1392

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                                                           |||   |||
Db       1393 ---------------------------------CCGCCCCATACCCCTTCGGTCTCCACC 1419

Qy        541 AspLysGluAspLeuLeuLys------GluLysThrAspAspThrSerGlyGluAspAsn 558
                       :::   :::         |||||||||||||||||||||||||||||||||
Db       1420 TCCGTGGGCAACTGCGTGGACTCAGGGGAGAAGACAGACGACACCTCAGGGGAGGACAAC 1479

Qy        559 AspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArg 578
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1480 GACGAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGC 1539

Qy        579 LysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrPro 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 AAAGGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCC 1599

Qy        599 GlnGlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGlu 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1600 CAGCAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAA 1659

Qy        619 GluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAla 638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1660 GAAATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCC 1719

Qy        639 ArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLys 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1720 CGGATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAG 1779

Qy        659 ArgGlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsn 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1780 AGGCAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAAC 1839

Qy        679 AlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProPro 698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1840 GCGCGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCC 1899

Qy        699 ValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetVal 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1900 GTGGTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTG 1959

Qy        719 GluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGly 738
              ||||||||||||||||||||||||||||||||||||
Db       1960 GAGGAGGCTGAAGCCTTACATGCCTCTGGGAATGAG------------------------ 1995

Qy        739 ProAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAla 758

Db       1995 ------------------------------------------------------------ 1995

Qy        759 AlaLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyPro 778

Db       1995 ------------------------------------------------------------ 1995

Qy        779 ProProGlyProProThrProProArgArgThrSerArgAlaProIleGluProThrPro 798

Db       1995 ------------------------------------------------------------ 1995

Qy        799 AlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaPro 818

Db       1995 ------------------------------------------------------------ 1995

Qy        819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGlu 838

Db       1995 ------------------------------------------------------------ 1995

Qy        839 GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858

Db       1995 ------------------------------------------------------------ 1995

Qy        859 GluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLys 878

Db       1995 ------------------------------------------------------------ 1995

Qy        879 AspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGly 898

Db       1995 ------------------------------------------------------------ 1995

Qy        899 GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
                                                           |||||||||||||||
Db       1996 ---------------------------------------------GACAGCGACTCCAGT 2010

Qy        919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2011 GCCACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTG 2070

Qy        939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2071 TCCCCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAG 2130

Qy        959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2131 AAGCCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATC------ 2184

Qy        979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAla 998

Db       2184 ------------------------------------------------------------ 2184

Qy        999 ProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSer 1018

Db       2184 ------------------------------------------------------------ 2184

Qy       1019 ProArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGlu 1038
                    |||:::   ||||||   ||||||
Db       2185 ------GGACGTGACCGCTCCTGGGCCCCACAA--------------------------- 2211

Qy       1039 AlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValPro 1058
                                            ||||||
Db       2212 ------------------------------TGGACA------------------------ 2217

Qy       1059 ProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAla 1078

Db       2218 ---------------------------------------------------GACATGTGC 2226

Qy       1079 ProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2227 TCCACAGGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGC 2286

Qy       1099 ProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeu 1118
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2287 CCACCCACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTC 2346

Qy       1119 GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSer 1138
              |||||||||||||||||||||||||||
Db       2347 GAGAGGCAAATAGGTGCCATCTCCCAA--------------------------------- 2373

Qy       1139 GluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPro 1158

Db       2373 ------------------------------------------------------------ 2373

Qy       1159 LysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAla 1178

Db       2373 ------------------------------------------------------------ 2373

Qy       1179 GlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThr 1198

Db       2373 ------------------------------------------------------------ 2373

Qy       1199 AlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPro 1218

Db       2373 ------------------------------------------------------------ 2373

Qy       1219 SerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeu 1238
                                                     |||||||||||||||||||||
Db       2374 ---------------------------------------GGCACGCCAGCTGACGTCCTG 2394

Qy       1239 TyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGly 1258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2395 TACAAGGGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGC 2454

Qy       1259 ArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeu 1278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2455 CGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTG 2514

Qy       1279 SerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSer 1298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2515 TCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTCCAAGGAGGACGGCAGAAGCAGCTCA 2574

Qy       1299 GlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVal 1318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2575 GGACCCCCCCATGAGACGGCCGCCCCCAAGCGCACCTATGACATGATGGAGGGCCGCGTG 2634

Qy       1319 GlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGlu 1338
              |||||||||||||||||||||||||||||||||
Db       2635 GGCAGAGCCATCTCCTCAGCCAGCATCGAAGGT--------------------------- 2667

Qy       1339 ArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358

Db       2667 ------------------------------------------------------------ 2667

Qy       1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378
                                                     |||   |||         |||
Db       2668 ---------------------------------GATAGCAGGGAGGAGACTTCATCTCTC 2694

Qy       1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
                       |||   |||                                 |||||||||
Db       2695 ---------GGTGCCCCC------------------------TGGTGGGCGGCCTACAAG 2721

Qy       1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2722 ACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGCCCATGAGGGCCTGGTGGCCACGGTG 2781

Qy       1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGlu 1438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2782 AAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCGCGAGGAGCTGCGGCACACGCCCGAG 2841

Qy       1439 LeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLys 1458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2842 CTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTCCATCACGCAGGGCACCCCGCTCAAG 2901

Qy       1459 TyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGly 1478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2902 TACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGC 2961

Qy       1479 SerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAla 1498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2962 AGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCA 3021

Qy       1499 LeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSer 1518
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3022 CTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCG 3081

Qy       1519 GlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGln 1538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3082 GGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCCGCGGCAG 3141

Qy       1539 SerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySer 1558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3142 AGCCCCCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCG 3201

Qy       1559 ProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLys 1578
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db       3202 CCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCCAGCAAG 3261

Qy       1579 AlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSer 1598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3262 GCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCGCACAGC 3321

Qy       1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3322 ACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGGGGCGTG 3381

Qy       1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3382 AGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCCATACCC 3441

Qy       1639 ArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsn 1658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3442 CGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAAC 3501

Qy       1659 ProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAla 1678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3502 CCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCG 3561

Qy       1679 LeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsn 1698
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3562 CTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAAC 3621

Qy       1699 ThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSer 1718
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3622 GCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCC 3681

Qy       1719 SerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValPro 1738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3682 TCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCA 3741

Qy       1739 HisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeu 1758
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3742 CACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTT 3801

Qy       1759 AlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSer 1778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3802 GCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCC 3861

Qy       1779 ProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArg 1798
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3862 CCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGA 3921

Qy       1799 AspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThr 1818
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3922 GACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACC 3981

Qy       1819 ThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSer 1838
              ||||||||||||||||||||||||||||||||||||
Db       3982 ACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTAGGGCATCAGAGCCCCCACCCCCC 4041
```

```
Qy        1839 GlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHis 1858
                                 |||                  ||||||      :::     |||
Db        4042 GCTCCGGGACTCCTTGTGGGCCGCAAGAGGCCTCCCCCTGCTGATGCCACTGACTGTCAC 4101

Qy        1859 GlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeu 1878
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4102 CAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTT 4161

Qy        1879 HisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeu 1898
               |||||||||||||||||||||||||||||||||||||||||||||    |||||||||||
Db        4162 CACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTG 4221

Qy        1899 ArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHis 1918
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4222 AGGTCCACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCAC 4281

Qy        1919 CysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeu 1938
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4282 TGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTG 4341

Qy        1939 ProLysGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAla 1958
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4342 CCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCC 4401

Qy        1959 PheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGly 1978
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4402 TTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGC 4461

Qy        1979 SerGluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrPro 1998
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4462 TCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCT 4521

Qy        1999 AlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAla 2018
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4522 GCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCC 4581

Qy        2019 SerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeu 2038
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4582 TCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTC 4641

Qy        2039 ArgSerLeu--------------------------------------------------- 2041
               |||||||||
Db        4642 CGTTCTCTGGGTAAGACCACCCTGACAGCGGCCACCTTCATAGACGCGATTATCATGCGT 4701

Qy        2041 ------------------------------------------------------------ 2041

Db        4702 CAAATTGCTCACGATAAAGGGGCGCGAGAAGGAGGTGCGCTGGCCAACGGCTCCCCTCGC 4761

Qy        2042 ---GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4762 GATGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGC 4821

Qy        2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4822 TCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGC 4881

Qy        2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4882 CACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCC 4941

Qy        2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4942 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 5001

Qy        2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5002 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 5061

Qy        2140 ------------------------------------------------------------ 2140

Db        5062 GTAACTACGCGTTCTCTGCTGCTGCTTGTCACCTTTGCACCTGGGGGCACCAGGCCTGGA 5121

Qy        2140 ------------------------------------------------------------ 2140

Db        5122 GAGGGGATGGGGAACCCCACAGCCCTTCTGTCCTGGCGGGGTGGCTGGGGGATCCAGGGC 5181

Qy        2141 ---------------------------GluValIleThrGlnAspTyrThrArgHisHis 2151
                                          |||||||||||||||||||||||||||||||||
Db        5182 ATGGCGCTGGGGGGATCCAGGGCGTGGGAGGTCATCACACAGGACTACACCCGGCACCAC 5241

Qy        2152 ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCys 2171
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5242 CCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGC 5301

Qy        2172 ProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGly 2191
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5302 CCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGT 5361

Qy        2192 AlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLys 2211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5362 GCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAG 5421

Qy        2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5422 ACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATG 5481

Qy        2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5482 ACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAG 5541

Qy        2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
               |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        5542 ACGGAGCCC---AGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTC 5598

Qy        2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5599 TTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAAC 5658

Qy        2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5659 AAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACG 5718

Qy        2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5719 GAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCG 5778

Qy        2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5779 GTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATG 5838

Qy        2352 GlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeu 2371
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5839 GGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCTCTG 5898

Qy        2372 AsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5899 AATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGTGAC 5958

Qy        2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5959 CACACACTCACCTCGCCAGGTGGCGGCGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGC 6018

Qy        2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6019 CGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCC 6078

Qy        2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6079 TCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAG 6138

Qy        2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6139 GACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTG 6198

Qy        2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6199 CAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTC 6258

Qy      2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6259 GCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAG 6318

Qy      2512 ThrLeuSerAspSerGlu 2517
             ||||||||||||||||||
Db      6319 ACACTCTCCGACAGCGAG 6336

RESULT 8
HSU37146
LOCUS       HSU37146                5989 bp    mRNA    linear   PRI 31-OCT-1995
DEFINITION  Human silencing mediator of retinoid and thyroid hormone action
            (SMRT) mRNA, complete cds.
ACCESSION   U37146
VERSION     U37146.1  GI:1045654
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5989)
  AUTHORS   Chen,J.D. and Evans,R.M.
  TITLE     A transcriptional co-repressor that interacts with nuclear hormone
            receptors
  JOURNAL   Nature 377 (6548), 454-457 (1995)
  MEDLINE   96008552
   PUBMED   7566127
REFERENCE   2  (bases 1 to 5989)
  AUTHORS   Chen,J.D. and Evans,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-1995) J. Don Chen, Gene Expression Lab, The Salk
            Institute, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .5989
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_line="HeLa"
     gene            1. .5989
                     /gene="SMRT"
     CDS             496. .4983
                     /gene="SMRT"
                     /note="transcriptional co-repressor"
                     /codon_start=1
                     /product="silencing mediator of retinoid and thyroid
                     hormone action"
                     /protein_id="AAC50236.1"
                     /db_xref="GI:1045655"
                     /translation="MEAWDAHPDKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKAS
                     PHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGA
                     ISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPE
                     SLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYK
                     GTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSS
                     GPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSIT
                     QGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGL
                     VATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHD
                     VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIV
                     PELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTST
                     PREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAA
                     AAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMA
                     QRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLP
                     TAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTT
                     VEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVL
                     HNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPV
                     LLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATI
                     ARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEG
                     VEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLP
                     ESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG
                     ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVS
                     PPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVK
                     SKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGL
                     EAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG
                     GKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAG
                     STPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDS
                     E"
     exon            4483. .4620
                     /gene="SMRT"
                     /note="alternatively spliced insert"
ORIGIN

Alignment Scores:
Pred. No.:              3.33e-151       Length:         5989
Score:                  7885.00         Matches:        1510
Percent Similarity:     94.34%          Conservative:   7
Best Local Similarity:  93.91%          Mismatches:     14
Query Match:            59.67%          Indels:         77
DB:                     9               Gaps:           7

US-09-522-753-5 (1-2517) x HSU37146 (1-5989)

Qy       984 ProProArgGluAspAlaAlaProThrLysProAlaProProAla--------------- 998
             |||||||||         :::|||   :::|||   |||||||||
Db       168 CCACCAAGA---------TCACCATATGAGCCACGACCCCCTGCATAGGAATAATCATAT 218

Qy       999 ProProProPro------------------------------------------------ 1002
             |||   ||||||
Db       219 CCGGAGCCACCCTGTGGTTCATAAGCCATCTGCCATTCTGATGGGCTCCATGTATCTATT 278

Qy      1003 ------------------GlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGly 1016
                                     |||:::|||         :::   :::   ||||||
Db       279 GCAGAGTCCCAAGTTTCATCAGCACTGAAACCAACCATGCCGTCGTAACGGTCTCCAGGT 338

Qy      1017 ---------------------------SerSerPro---------------------Arg 1020
                                        |||   |||                     |||
Db       339 CTCCCTCTTCTGTCATAGGCCATGAGGTCTCCCCCTCTAGGTGGTGGTGGAGGAAGA 398

Qy      1021 GlyLysSerArgSer---------ProAlaProPro-Ala-------------------- 1030
             |||:::   |||:::         |||   ||||||
Db       399 GGAAGATTCCGAGCTCTGCTACCACCCCGGCCGCCTCGTCCGGGAGGAGGGGGAGGTGGT 458

Qy      1030 ------------------------------------------------------------ 1030

Db       459 CCTCGACGAGGGCTCATATCATCATAATCTCTTCTAGATGGAGGCATGGGACGCCCACCC 518

Qy      1031 -AspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpTh 1050
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       519 CGACAAGGAGGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGAC 578

Qy      1050 rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       579 TTCCGGCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCC 638

Qy      1070 oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       639 GGACCCCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGA 698

Qy      1090 pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       699 CACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTC 758

Qy      1110 rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       759 CTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTC 818

Qy      1130 rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       819 GGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCAT 878

Qy      1150 tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       879 GGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGA 938

Qy      1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 GCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCACAGCCCA 998

Qy       1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 GGAGGCGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAA 1058

Qy       1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 AGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCAC 1118

Qy       1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1119 CCACGGCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGA 1178

Qy       1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 CAGCCCGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTA 1238

Qy       1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 CGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTC 1298

Qy       1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1299 CAAGGAGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCAC 1358

Qy       1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 CTATGACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCT 1418

Qy       1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1419 CATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCA 1478

Qy       1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1479 CATCCGCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTA 1538

Qy       1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1539 CCTGCGTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTC 1598

Qy       1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1599 ACGGGACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGC 1658

Qy       1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1659 CCATGAGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCG 1718

Qy       1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySe 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1719 CGAGGAGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTC 1778

Qy       1450 rIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLy 1470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1779 CATCACGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAA 1838

Qy       1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLe 1490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1839 GCACGACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCT 1898

Qy       1490 uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSe 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1899 GGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAG 1958

Qy       1510 rArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa 1530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1959 CCGGCCAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGT 2018

Qy       1530 lProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPh 1550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2019 GCCTGAGCTGGGTAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCACGGGGCACCCTT 2078

Qy       1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl 1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2079 TGCCGGCCACCTCCCACGAGGTTCGCCCGTGACCATGCGGGAGCCCACGCCGCGCCTGCA 2138

Qy       1570 nGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2139 GGAGGGCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCG 2198

Qy       1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2199 TGAGATCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCC 2258

Qy       1610 oTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLe 1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2259 CTATGAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCT 2318

Qy       1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTy 1650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2319 GGCCTTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTA 2378

Qy       1650 rLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIl 1670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2379 CCTGCCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCAT 2438

Qy       1670 eArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIl 1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2439 CCGCGGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACAT 2498

Qy       1690 eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2499 CACCTCGCAGCAGATGCACCACAACACGGCCACCGCCATGGCCCAGCGAGCTGATATGCT 2558

Qy       1710 uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGl 1730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2559 GAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGG 2618

Qy       1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2619 CATCATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCAC 2678

Qy       1750 rProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSe 1770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2679 CCCAGCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAG 2738

Qy       1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh 1790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2739 CCGCCACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCAC 2798

Qy       1790 rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr 1810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2799 CACGTCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCG 2858

Qy       1810 gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGl 1830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2859 GGAAAAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGG 2918

Qy       1830 yThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerAr 1850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2919 TACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCG 2978

Qy       1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2979 CCCCGCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGC 3038

Qy       1870 aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa 1890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3039 CCTCCAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGT 3098

Qy       1890 lGluProSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAl 1910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3099 GGAGCCCAGCAAGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGC 3158
```

```
Qy        1910 aAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPr 1930
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3159 TGCCACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCC 3218

Qy        1930 oThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluAr 1950
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3219 TACCCTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCG 3278

Qy        1950 gProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGl 1970
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3279 GCCCCGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGA 3338

Qy        1970 uProAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGl 1990
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3339 GCCCGCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGG 3398

Qy        1990 yHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAs 2010
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3399 CCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGA 3458

Qy        2010 pProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysPr 2030
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3459 CCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACC 3518

Qy        2030 oPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerPr 2050
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3519 CTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCC 3578

Qy        2050 oGluGlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLe 2070
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3579 CGAAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCT 3638

Qy        2070 uProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGl 2090
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3639 CCCCAAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCA 3698

Qy        2090 nProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuPr 2110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3699 GCCAGGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCC 3758

Qy        2110 oGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGl 2130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3759 TGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCA 3818

Qy        2130 nArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHi 2150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3819 GCGGGTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCA 3878

Qy        2150 sHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSe 2170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3879 CCACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAG 3938

Qy        2170 rCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHi 2190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3939 CTGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCA 3998

Qy        2190 sGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAs 2210
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3999 TGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAA 4058

Qy        2210 nLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGl 2230
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4059 CAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGG 4118

Qy        2230 yMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGl 2250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4119 CATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGA 4178

Qy        2250 uGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAl 2270
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4179 ACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGC 4238
```

```
Qy        2270 aPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIl 2290
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4239 CTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGAT 4298

Qy        2290 eAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGl 2310
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4299 CAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGG 4358

Qy        2310 yThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGl 2330
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4359 GACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCA 4418

Qy        2330 nAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLe 2350
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4419 GGCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACT 4478

Qy        2350 uMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnPr 2370
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4479 CATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCC 4538

Qy        2370 oLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSe 2390
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4539 TCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAG 4598

Qy        2390 rAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSe 2410
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4599 TGACCACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAG 4658

Qy        2410 rSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerVa 2430
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4659 CAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGT 4718

Qy        2430 lSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTr 2450
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4719 CTCCTCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTG 4778

Qy        2450 pGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetAr 2470
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4779 GGAGGACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCG 4838

Qy        2470 gLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyPr 2490
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4839 GCTGCAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCC 4898

Qy        2490 oLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTy 2510
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4899 CCTCGCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTA 4958

Qy        2510 rGluThrLeuSerAspSerGlu 2517
               ||||||||||||||||||||||
Db        4959 CGAGACACTCTCCGACAGCGAG 4980

RESULT 9
S83390
LOCUS       S83390                  2930 bp    mRNA    linear   PRI 28-MAR-1997
DEFINITION  T3 receptor-associating cofactor-1 [human, fetal liver, mRNA, 2930
            nt].
ACCESSION   S83390
VERSION     S83390.1  GI:1911769
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2930)
  AUTHORS   Sande,S. and Privalsky,M.L.
  TITLE     Identification of TRACs (T3 receptor-associating cofactors), a
            family of cofactors that associate with, and modulate the activity
            of, nuclear hormone receptors
  JOURNAL   Mol. Endocrinol. 10 (7), 813-825 (1996)
  MEDLINE   96408715
   PUBMED   8813722
  REMARK    GenBank staff at the National Library of Medicine created this
```

```
            entry [NCBI gibbsq 179975] from the original journal article.
FEATURES             Location/Qualifiers
     source          1. .2930
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1. .2930
                     /gene="T3 receptor-associating cofactor-1, TRAC-1"
     CDS             456. .2765
                     /gene="T3 receptor-associating cofactor-1, TRAC-1"
                     /note="Authors report detecting several TRAC  alternative
                     transcripts; TRAC-1"
                     /codon_start=1
                     /product="T3 receptor-associating cofactor-1"
                     /protein_id="AAB50847.1"
                     /db_xref="GI:1911770"
                     /translation="MAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPP
                     TPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRER
                     DRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSP
                     ISPRTQDALQQRPSVLHNTGMKGIITAVEPSTPTVLRSTSTSSPVRPAATFPPATHCP
                     LGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKG
                     SEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQEL
                     ELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPV
                     KLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHP
                     QQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPN
                     KTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP
                     PAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMT
                     YRSQAVQEHASTNMGLEAIIRKALMGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSV
                     SSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGS
                     GPLAGAHHAWDEEPKPLLCSQYETLSDSE"
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-85        Length:         2930
Score:                  4612.00         Matches:        873
Percent Similarity:     94.38%          Conservative:   1
Best Local Similarity:  94.28%          Mismatches:     6
Query Match:            34.90%          Indels:         46
DB:                     9               Gaps:           1

US-09-522-753-5 (1-2517) x S83390 (1-2930)

Qy       1592 IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerProTyr 1611
              :::         ||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CTTGGCTCACCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTAT 182

Qy       1612 GluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAla 1631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCC 242

Qy       1632 PheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeu 1651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTG 302

Qy       1652 ProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArg 1671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGC 362

Qy       1672 GlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThr 1691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 GGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACC 422

Qy       1692 SerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArg 1711
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        423 TCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGG 482

Qy       1712 GlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIle 1731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 GGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATC 542

Qy       1732 IleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrPro 1751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCA 602

Qy       1752 AlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg 1771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 GCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGC 662

Qy       1772 HisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThr 1791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 CACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACG 722

Qy       1792 SerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGlu 1811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 TCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAA 782

Qy       1812 LysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThr 1831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACA 842

Qy       1832 GluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgPro 1851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 GAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCC 902

Qy       1852 AlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeu 1871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTC 962

Qy       1872 GlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGlu 1891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAG 1022

Qy       1892 ProSerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAla 1911
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CCCAGCACGCCCACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCAGCTGCC 1082

Qy       1912 ThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThr 1931
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 ACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACC 1142

Qy       1932 LeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgPro 1951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 CTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCC 1202

Qy       1952 ArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluPro 1971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 CGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCC 1262

Qy       1972 AlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHis 1991
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCAC 1322

Qy       1992 AlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspPro 2011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1323 GCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCG 1382

Qy       2012 ProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPhe 2031
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1383 CCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTT 1442

Qy       2032 SerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGlu 2051
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 TCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAA 1502

Qy       2052 GlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuPro 2071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 GGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCC 1562

Qy       2072 LysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnPro 2091
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1563 AAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCA 1622

Qy       2092 GlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGlu 2111
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 GGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCGCTGCCTGAG 1682
```

```
Qy        2112 SerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArg 2131
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1683 AGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAAAGGTCACCAGCGG 1742

Qy        2132 ValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHis 2151
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1743 GTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTACACCCGGCACCAC 1802

Qy        2152 ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCys 2171
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1803 CCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGCTGC 1862

Qy        2172 ProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHisGly 2191
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1863 CCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCATGGT 1922

Qy        2192 AlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLys 2211
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1923 GCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAG 1982

Qy        2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1983 ACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGCATG 2042

Qy        2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2043 ACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAACAG 2102

Qy        2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2103 ACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTC 2162

Qy        2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2163 TTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATCAAC 2222

Qy        2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2223 AAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGGACG 2282

Qy        2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2283 GAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAGGCG 2342

Qy        2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2343 GTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTCATG 2402

Qy        2352 GlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeu 2371

Db        2402 ------------------------------------------------------------ 2402

Qy        2372 AsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391

Db        2402 ------------------------------------------------------------ 2402

Qy        2392 HisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
                                 ||||||||||||||||||||||||||||||||||||||||||
Db        2403 ------------------GGTGGCGGCGGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGC 2444

Qy        2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2445 CGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCCACCCTCTGTCTCC 2504

Qy        2432 SerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2505 TCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAACCGCGTGTGGGAG 2564

Qy        2452 AspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeu 2471
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2565 GACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCTGATCATGCGGCTG 2624

Qy        2472 GlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu 2491
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2625 CAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGGCAGCGGGCCCCTC 2684

Qy        2492 AlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2685 GCTGGCGCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTGCTCGCAGTACGAG 2744

Qy        2512 ThrLeuSerAspSerGlu 2517
               ||||||||||||||||||
Db        2745 ACACTCTCCGACAGCGAG 2762

RESULT 10
AB028970
LOCUS       AB028970                7949 bp    mRNA    linear   PRI 02-APR-2003
DEFINITION  Homo sapiens mRNA for KIAA1047 protein, partial cds.
ACCESSION   AB028970
VERSION     AB028970.2  GI:29421187
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kikuno,R., Nagase,T., Ishikawa,K., Hirosawa,M., Miyajima,N.,
            Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XIV. The complete sequences of 100 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 6 (3), 197-205 (1999)
  MEDLINE   99397452
   PUBMED   10470851
REFERENCE   2  (bases 1 to 7949)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
COMMENT     On Mar 31, 2003 this sequence version replaced gi:5689430.
FEATURES             Location/Qualifiers
     source          1. .7949
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /clone="hh01221s1"
                     /tissue_type="brain"
                     /note="vector: pBC SK plus"
     gene            1. .7949
                     /gene="KIAA1047"
     CDS             <12. .7058
                     /gene="KIAA1047"
                     /note="Start codon is not identified.
                     This sequence was obtained by subcloning of the DNA
                     fragments derived from two cDNA clones (1 - 1691 was
                     derived from hh01221 and 1692 - 7949 was derived from
                     hg02672)."
                     /codon_start=1
                     /product="KIAA1047 protein"
                     /protein_id="BAA82999.2"
                     /db_xref="GI:29421188"
                     /translation="SHSLVRTSLLIMSSSGYPPNQGAFSTEQSRYPPHSVQYTFPNTR
                     HQQEFAVPDYRSSHLEVSQASQLLQQQQQQLRRRPSLLSEFHPGSDRPQERRTSYEP
                     FHPGPSPVDHDSLESKRPRLEQVSDSHFQRVSAAVLPLVHPLPEGLRASADAKKDPAF
                     GGKHEAPSSPISGQPCGDDQNASPSKLSKEELIQSMDRVDREIAKVEQQILKLKKKQQ
                     QLEEEAAKPPEPEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKIFEGLGPKVELPL
                     YNQPSDTKVYHENIKTNQVMRKKLILFFKRRNHARKQREQKICQRYDQLMEAWEKKVD
                     RIENNPRRKAKESKTREYYEKQFPEIRKQREQQERFQRVGQRGAGLSATIARSEHEIS
                     EIIDGLSEQENNEKQMRQLSVIPPMMFDAEQRRVKFINMNGLMEDPMKVYKDRQFMNV
                     WTDHEKEIFKDKFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKALVRRNYGK
                     RRGRNQQIARPSQEEKVEEKEEDKAEKTEKKEEEKKDEEEKDEKEDSKENTKEKDKID
                     GTAEETEEREQATPRGRKTANSQGRRKGRITRSMTNEAAAASAAAAAATEEPPPPLPP
                     PPEPISTEPVETSRWTEEEMEVAKKGLVEHGRNWAAIAKMVGTKSEAQCKNFYFNYKR
```

```
                    RHNLDNLLQQHKQKTSRKPREERDVSQCESVASTVSAQEDEDIEASNEEENPEDSEGA
                    ENSSDTESAPSPSPVEAVKPSEDSPENATSRGNTEPAVELEPTTETAPSTSPSLAVPS
                    TKPAEDESVETQVNDSISAETAEQMDVDQQEHSAEEGSVCDPPPATKADSVDVEVRVP
                    ENHASKVEGDNTKERDLDRASEKVEPRDEDLVVAQQINAQRPEPQSDNDSSATCSADE
                    DVDGEPERQRMFPMDSKPSLLNPTGSILVSSPLKPNPLDLPQLQHRAAVIPPMVSCTP
                    CNIPIGTPVSGYALYQRHIKAMHESALLEEQRQRQEQIDLECRSSTSPCGTSKSPNRE
                    WEVLQPAPHQVITNLPEGVRLPTTRPTRPPPPLIPSSKTTVASEKPSFIMGGSISQGT
                    PGTYLTSHNQASYTQETPKPSVGSISLGLPRQQESAKSATLPYIKQEEFSPRSQNSQP
                    EGLLVRAQHEGVVRGTAGAIQEGSITRGTPTSKISVESIPSLRGSITQGTPALPQTGI
                    PTEALVKGSISRMPIEDSSPEKGREEAASKGHVIYEGKSGHILSYDNIKNAREGTRSP
                    RTAHEISLKRSYESVEGNIKQGMSMRESPVSAPLEGLICRALPRGSPHSDLKERTVLS
                    GSIMQGTPRATTESFEDGLKYPKQIKRESPPIRAFEGAITKGKPYDGITTIKEMGRSI
                    HEIPRQDILTQESRKTPEVVQSTRPIIEGSISQGTPIKFDNNSGQSAIKHNVKSLITG
                    PSKLSRGMPPLEIVPENIKVVERGKYEDVKAGETVRSRHTSVVSSGPSVLRSTLHEAP
                    KAQLSPGIYDDTSARRTPVSYQNTMSRGSPMMNRTSDVTISSNKSTNHERKSTLTPTQ
                    RESIPAKSPVPGVDPVVSHSPFDPHHRGSTAGEVYRSHLPTHLDPAMPFHRALDPAAA
                    AYLFQRQLSPTPGYPSQYQLYAMENTRQTILNDYITSQQMQVNLRPDVARGLSPREQP
                    LGLPYPATRGIIDLTNMPPTILVPHPGGTSTPPMDRITYIPGTQITFPPRPYNSASMS
                    PGHPTHLAAAASAERERERERERERKERERERIAAASSDLYLRPGSEQPGRPGSHGYVRSP
                    SPSVRTQETMLQQRPSVFQGTNGTSVITPLDPTAQLRIMPLPAGGPSISQGLPASRYN
                    TAADALAALVDAAASAPQMDVSKTKEISSHRYETPSDAIEVISPASSPAPPQEKLQTY
                    QPEVVKANQAENDPTRQYEGPLHHYRPQQESPSPQQQLPPSSQAEGMGQVPRTHRLIT
                    LADHICQIITQDFARNQVSSQTPQQPPTSTFQNSPSALVSTPVRTKTSNRYSPESQAQ
                    SVHHQRPGSRVSPENLVDKSRGSRPGKSPERSHVSSEPYEPISPPQVPVVHEKQDSLL
                    LLSQRGAEPAEQRNDARSPGSISYLPSFFTKLENTSPMVKSKKQEIFRKLNSSGGGDS
                    DMAAAQPGTEIFNLPAVTTSGSVSSRGHSFADPASNLGLEDIIRKALMGSFDDKVEDH
                    GVVMSQPMGVVPGTANTSVVTSGETRREEGDPSPHSGGVCKPKLISKSNSRKSKSPIP
                    GQGYLGTERPSSVSSVHSEGDYHRQTPGWAWEDRPSSTGSTQFPYNPLTMRMLSSTPP
                    TPIACAPSAVNQAAPHQQNRIWEREPAPLLSAQYETLSDSDD"
ORIGIN

Alignment Scores:
Pred. No.:              1.24e-76       Length:          7949
Score:                  4194.00        Matches:         1097
Percent Similarity:     53.92%         Conservative:    340
Best Local Similarity:  41.16%         Mismatches:      720
Query Match:            31.74%         Indels:          508
DB:                     9              Gaps:            108

US-09-522-753-5 (1-2517) x AB028970 (1-7949)

Qy          16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
               :::   ||||||||||||||||||:::   |||               |||
Db          93 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 152

Qy          36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
               :::      :::   :::|||:::      |||               ||||||   |||
Db         153 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 212

Qy          54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                 |||      ||||||||||||||||||||||||||||||
Db         213 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 272

Qy          74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
               ||||||::::::|||   ||||||         |||               :::   |||
Db         273 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 326

Qy          94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                              :::   :::   :::||||||||||||||||||||||||   :::
Db         327 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 386

Qy         114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                  |||                                 ||||||:::   |||   |||
Db         387 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 446

Qy         123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                           |||      |||   |||   |||      ||||||   :::
Db         447 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 494

Qy         143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
               ||||||   |||   |||   ||||||   |||                  :::
Db         495 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 554

Qy         162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                     |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db         555 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 614

Qy         182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
               ||||||      |||||||||||||||   ||||||||||||||||||||||||||||||
Db         615 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 674

Qy         202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db         675 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 734

Qy         222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
               ||||||||||||:::|||||||||||||||||||||||||||||||||||||||   |||
Db         735 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAGAAAGCAGAAGAAGCT 794

Qy         242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
               |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db         795 CATAAAATTTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 854

Qy         262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIle 281
               ||||||:::   ||||||||||||||||||   ||||||   ||||||||||||||||||
Db         855 GATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATT 914

Qy         282 LeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArg 301
               |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db         915 TTATTTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAGCGT 974

Qy         302 TyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArg 321
               |||||||||||||||||||||   ||||||||||||:::|||||||||||||||||||||
Db         975 TATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGG 1034

Qy         322 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArg 341
               |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        1035 AGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGA 1094

Qy         342 LysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeu 361
               ||||||||||||   |||||||||   |||   ||||||||||||||||||:::||||||
Db        1095 AAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGGCAGAGGGGAGCTGGTCTT 1151

Qy         362 SerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGlu 381
               |||   :::   ||||||||||||||||||:::|||||||||||||||||||||||||||
Db        1152 TCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGGCTCTCTGAG 1211

Qy         382 GlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAsp 401
               |||||||||   |||||||||||||||||||||:::|||||||||||||||::::::|||
Db        1212 CAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGATTCCACCTATGATGTTTGAT 1271

Qy         402 AlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLys 421
               |||:::|||:::|||:::|||||||||||||||||||||||||||   ||||||||||||
Db        1272 GCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCTTATGGAGGACCCTATGAAA 1331

Qy         422 ValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArg 441
               ||||||||||||||||||   ||||||:::|||::::::   |||||||||   |||:::
Db        1332 GTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAG 1391

Qy         442 GluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLys 461
               :::||||||:::|||||||||||||||||||||||||||||||||:::||||||||||||
Db        1392 GACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCATCATACTTGGAGAGGAAG 1451

Qy         462 ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSer 481
               :::|||   :::|||||||||||||||||||||||||||||||||||||||||||||:::
Db        1452 AGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAAAATGAGAATTATAAAGCC 1511

Qy         482 LeuValArgArgSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
               ||||||||||||:::|||   :::|||||||||::::::||||||      :::
Db        1512 CTCGTCAGAAGGAATTATGGGAAACGCAGAGGCAGAAACCAGCAAATTGCTCGACCCTCG 1571

Qy         501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
               |||:::::::::   :::::::::::::::         ::::::   :::   ||||||
Db        1572 CAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAAGCAGAAAAAACAGAAAAA 1625

Qy         521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
```

```
                |||:::||||||:::   |||   ||||||:::|||::::::         ||||||
Db      1626 AAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAAGAAGACTCCAAAGAAAAT 1685

Qy       541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||         |||:::|||   |||   |||:::   |||:::   :::|||
Db      1686 ACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA---GAAGAAACTGAGGAA 1733

Qy       561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
             :::|||            :::|||||||||||||||||||||||||||||||||||||||
Db      1734 AGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGTCAGGGCCGCCGTAAGGGC 1793

Qy       581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
             ||||||||||||||||||   |||||||||   :::      |||
Db      1794 CGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGTGCTGCAGCCGCAGCGGCT 1853

Qy       601 Ser---------------------------AlaGluLeuAlaSerMetGluLeuAsn 610
             :::                                 |||      |||   |||
Db      1854 ACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCCATTTCTACAGAGCCTGTG 1913

Qy       611 GluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHis 630
             |||:::|||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db      1914 GAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGTCTAGTAGAACAT 1973

Qy       631 GlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLys 650
             ||||||||||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||
Db      1974 GGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAAAGTGAAGCTCAATGTAAA 2033

Qy       651 AsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLys 670
             |||||||||||||||||||||:::|||   |||||||||   :::|||||||||||||||
Db      2034 AACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAACCTCTTACAGCAGCATAAA 2093

Qy       671 LeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAla 690
                |||      ::::::            |||         :::      :::   |||
Db      2094 CAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCTCAATGTGAAAGTGTCGCT 2153

Qy       691 SerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyVal 710
             |||      :::|||               |||||||||::::::|||||||||
Db      2154 TCCACTGTTTCTGCT------------CAGGAGGATGAAGATATTGAAGCCTCC------ 2195

Qy       711 SerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGlu 730
                   ||||||||||||      |||::::::|||
Db      2196 ------AATGAAGAAGAAAATCCAGAAGACAGCGAAGGT--------------------- 2228

Qy       731 ValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSer 750
                                                    |||||||||||||||||||||
Db      2229 ---------------------------------GCAGAAAATAGTTCTGATACAGAAAGT 2255

Qy       751 IleProSerProHis---ThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysPro 769
                |||||||||         ||||||   |||
Db      2256 GCTCCTTCTCCTTCACCAGTTGAAGCTGTCAAG--------------------------- 2288

Qy       770 ProAlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThr 789
                                           |||         :::|||         |||
Db      2289 ------------------------------CCCAGCGAGGACAGTCCTGAAAATGCTACT 2318

Qy       790 SerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProPro 809
             ||||||         ||||||            |||            ||||||
Db      2319 TCTCGAGGAAACACAGAACCTGCGGTTGAGCTTGAG-----------CCCACCACGGAA 2366

Qy       810 ProAlaPro---ProSerProSer---AlaProProProValValProLysGluGluLys 827
                ||||||      |||||||||   |||   |||         |||   |||::::::
Db      2367 ACTGCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACAAAACCAGCTGAAGATGAA 2426

Qy       828 GluGluGluThrAlaAlaAlaProProVal------------------------------ 837
                   ||||||               :::
Db      2427 AGTGTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACAGCAGAGCAGATGGATGTA 2486

Qy       838 ---------------------GluGluGlyGluGluGlnLysProProAlaAlaGluGlu 850
                                  |||||||||            ||||||   |||   :::
Db      2487 GATCAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGATCCCCCACCCGCTACCAAA 2546

Qy       851 LeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAla 870
                      |||:::         |||      |||                  :::
Db      2547 ------GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAAAACCATGCATCTAAAGTT 2600

Qy       871 GluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAla 890
             |||            |||   :::|||   :::   |||
Db      2601 GAAGGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC--------------------- 2639

Qy       891 LeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGly 910
                   :::||||||   |||                        |||:::
Db      2640 ------AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTGGTAGCTCAGCAAATAAAT 2693

Qy       911 Ala------------ProGln---AspSerAspSerSerAlaThrCysSerAlaAspGlu 925
             |||            ||||||   |||:::||||||||||||||||||||||||||||||
Db      2694 GCCCAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCCACGTGCAGCGCTGATGAG 2753

Qy       926 ValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro---------ArgPro 942
                   :::   :::|||      :::   :::   :::   |||         :::|||
Db      2754 ------GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTTCCTATGGACTCAAAGCCT 2807

Qy       943 SerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------Pro 960
             |||||||||   |||||||||               :::||||||   |||      |||
Db      2808 TCACTGTTAAACCCCACTGGATCTATACTCGTC---TCATCTCCGTTAAAACCAAATCCA 2864

Qy       961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIle------------ 976
             |||||||||   ||||||:::   |||||||||   |||||||||:::
Db      2865 CTGGATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCACCAATGGTATCCTGCACC 2924

Qy       977 ------------------------------------------------------GlnVal 978
                                                                      :::
Db      2925 CCATGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCTCTCTACCAGCGACACATT 2984

Qy       979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAla 998
                   :::||||||            :::
Db      2985 AAAGCAATGCATGAGTCAGCACTCCTGGAG------------------------------ 3014

Qy       999 ProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSer 1018
                         :::      |||   :::   |||         :::      :::|||
Db      3015 ------GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAATGTAGAAGTTCTACAAGT 3068

Qy      1019 ProArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGlu 1038
             |||   |||   |||:::||||||            ::::::|||
Db      3069 CCATGTGGCACATCCAAGAGTCCA------------AACAGAGAG--------------- 3101

Qy      1039 AlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValPro 1058
                                           |||
Db      3102 ------------------------------TGG--------------------------- 3104

Qy      1059 ProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAla 1078
                   ||||||::::::   :::||||||                     :::
Db      3105 ------GAAGTCCTTCAGCCTGCTCCACAT------------CAAGTGATAACTAATCTC 3146

Qy      1079 ProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
             |||   |||      ||||||               |||   ||||||      |||
Db      3147 CCTGAAGGCGTTCGGCTTCCG---------------ACAACTCGACCAACCAGGCCACCG 3191

Qy      1099 ProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeu 1118
             ||||||   |||   :::            ::::::|||   |||   ||||||   :::
Db      3192 CCCCCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAAAAA---CCATCTTTATA 3248

Qy      1119 GluArgGlnIleGlyAlaIleSerGlnGly---------------MetSerValGlnLeu 1133
                         |||:::||||||||||||               ::::::
Db      3249 ------ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACTTATTTGACTTCTCATAAT 3302

Qy      1134 HisValProTyrSerGluHisAlaLysAlaPro---ValGlyProValThrMetGlyLeu 1152
                      |||::::::            |||   ||||||   :::::::::||||||
Db      3303 CAGGCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTGGGATCTATCTCTCTTGGACTG 3362

Qy      1153 ProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeu 1172
             |||         :::      |||   |||            :::|||||||||:::
Db      3363 CCACGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTACATCAAGCAGGAAGAATTT 3422

Qy      1173 SerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAla 1192
             |||||||||   |||         ||||||   |||   |||      ||||||
Db      3423 TCTCCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC---AGGGCCCAACATGAA 3479
```

```
Qy        1193 SerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIle 1212
                  |||:::||||||||||||   |||::::::      ||||||||||||:::|||
Db        3480 GGTGTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGAAGTATAACTCGGGGAACT 3536

Qy        1213 ProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGly 1232
               |||::::::::::::      :::|||      :::   |||||||||||||||   |||
Db        3537 CCAACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGGGGCTCTATCACTCAGGGC 3596

Qy        1233 ThrProAla------------------------AspValLeuTyrLysGlyThrIleThr 1244
               |||||||||                        :::   |||   ||||||:::|||:::
Db        3597 ACCCCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTGGTGAAGGGGTCCATTTCG 3656

Qy        1245 ArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuPro 1264
               |||:::      |||||||||            ::::::|||||||||::::::
Db        3657 AGAATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGCAGAGAGGAAGCTGCATCC 3710

Qy        1265 LysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMet 1284
               |||||||||||||||||||||||||||   ||||||:::|||||||||:::
Db        3711 AAAGGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTGTCATATGATAATATTAAG 3770

Qy        1285 SerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThr 1304
               :::                     :::|||      |||            ||||||
Db        3771 AATGCC---------------CGAGAAGGGACTAGGAGTCCAAGAACAGCTCATGAAATC 3815

Qy        1305 AlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer--- 1323
               :::      ||||||:::|||:::   :::||||||   :::   :::   :::|||
Db        3816 AGTTTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATAAAGCAAGGATGTCAATG 3872

Qy        1324 ---------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGlu 1338
                              ||||||   :::|||||||||:::   ||||||:::|||
Db        3873 AGGGAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGCCGAGCATTACCC------ 3926

Qy        1339 ArgHisSerProHis---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGln 1357
               |||   |||||||||      |||||||||:::      :::   |||||||||   |||
Db        3927 AGGGGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTGTCTGGCTCCATAATGCAG 3986

Qy        1358 GlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeu 1377
               |||   ||||||:::      |||:::   ||||||   |||   :::      |||
Db        3987 GGGACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT---AAATATCCCAAACAA 4043

Qy        1378 LeuLysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyr 1397
               :::|||||||||                  ||||||   |||            |||
Db        4044 ATTAAAAGGGAAAGT---------------CCTCCCATACGAGCATTTGAAGGTGCCATT 4088

Qy        1398 LysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThr 1417
                                       |||   ||||||   ::::::|||   :::   |||
Db        4089 ---------------------ACCAAAGGAAAACCA---TATGATGGC---ATCACCACC 4121

Qy        1418 ValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu------------ 1433
               :::||||||   |||||||||||||||||||||||||||:::::::::
Db        4122 ATCAAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAAGATATTTTAACTCAGGAA 4181

Qy        1434 ---ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThr 1452
                  |||   |||||||||:::      :::   ||||||:::   ||||||||||||:::
Db        4182 AGTCGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATAATTGAGGGTTCCATTTCC 4241

Qy        1453 GlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAsp 1472
               ||||||||||||:::|||:::|||         |||         :::   ||||||:::
Db        4242 CAGGGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAATCTGCCATCAAACACAAT 4298

Qy        1473 ValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVal 1492
               |||:::|||||||||      |||   :::            :::   ||||||::::::
Db        4299 GTCAAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGAATGCCTCCGCTGGAAATT 4358

Qy        1493 MetAlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu---------------Glu 1506
               :::   :::      :::   :::||||||      ||||||               |||
Db        4359 GTGCCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAG 4418

Qy        1507 SerLeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAla 1526
               :::::::::||||||      :::      |||||||||   |||:::   |||
Db        4419 ACCGTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCCTCCGTTCTTAGGTCCACA 4478

Qy        1527 ProValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHis 1546
                        :::   |||      |||   :::   ||||||      |||:::|||
Db        4479 ---------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCTGGGATTTATGATGACACC 4529

Qy        1547 GlyAla---------------ProPheAlaGlyHisLeuProArgGlySerProValThr 1561
                  |||                  :::         :::   ||||||||||||:::
Db        4530 AGTGCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCCAGAGGCTCACCCATG--- 4586

Qy        1562 MetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAla---Ser 1580
                                 |||      :::   ::::::||||||:::|||:::   :::
Db        4587 ------------ATGAACAGAACTTCTGATGTTACAATTTCTTCTAACAAGTCTACCAAT 4634

Qy        1581 GlnAspArgLysLeuThrSerThrProArgGlu------------IleAlaLysSerPro 1596
                  :::||||||   |||   ||||||   :::               ||||||||||||
Db        4635 CATGAAAGGAAATCGACACTGACCCCTACCCAGAGGGAAAGTATCCCAGCGAAGTCTCCA 4694

Qy        1597 HisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArg 1616
                        |||               |||      ||||||::::::         |||
Db        4695 GTGCCTGGGGTGGACCCTGTCGTGAGCCAC------AGTCCGTTTGATCCCCATCACAGA 4748

Qy        1617 GlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSer 1636
               |||   :::      ::::::||||||||||||:::|||         ||||||   :::
Db        4749 GGCAGCACTGCAGGCGAGGTTTATCGGAGCCACCTGCCCACGCACTTGGATCCA---GCC 4805

Qy        1637 IleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHisLeu 1655
               :::|||            ||||||   |||||||||||||||         |||   |||
Db        4806 ATGCCTTTTCACAGGGCTTTGGATCCTGCAGCGGCTGCTTACCTGTTTCAGAGACAGCTT 4865

Qy        1656 AlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAsp 1675
               :::|||   |||   ||||||      |||      |||
Db        4866 TCACCAACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC--------------------- 4904

Qy        1676 ThrAlaAlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGln 1694
                     |||:::||||||   ||||||||||||:::||||||||||||||||||||||||
Db        4905 ------GCAATGGAGAACACAAGACAGACAATCTTAAATGATTACATTACCTCACAACAG 4958

Qy        1695 MetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSer 1714
               |||      |||                     |||   |||:::   ||||||||||||
Db        4959 ATGCAAGTGAACTTG------------------CGTCCAGATGTGGCCAGAGGACTCTCC 5000

Qy        1715 ProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeu 1734
               |||||||||      |||   |||   |||   |||      ||||||||||||||||||
Db        5001 CCAAGAGAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGAGGAATCATTGACCTG 5057

Qy        1735 SerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAla 1754
               :::   :::|||         :::|||||||||         ||||||
Db        5058 ACCAATATGCCTCCA---ACAATTTTAGTGCCTCATCCAGGGGGAACAAGCACTCCTCCC 5114

Qy        1755 MetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg---HisSer 1773
               |||||||||:::   |||:::|||               |||      |||   ::::::
Db        5115 ATGGACAGAATCACTTATATTCCTGGTACACAGATTACTTTCCCTCCCAGGCCGTACAAC 5174

Qy        1774 SerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSer 1793
               |||:::   :::|||||||||   ||||||||||||                  :::|||
Db        5175 TCTGCTTCCATGTCTCCAGGACACCCAACACACCTT---------GCAGCTGCTGCAAGT 5225

Qy        1794 SerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
               :::||||||||||||:::|||:::||||||::::::|||:::|||||||||
Db        5226 GCTGAGAGGGAACGGGAACGGGAGCGGGAGAAGGAGCGGGAGCGGGAACGG--------- 5276

Qy        1814 IleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGln 1833
               |||      :::::::::   :::                  |||||||||:::||||||
Db        5277 ATTGCTGCAGCTTCCTCCGACCTC------------TACCTGCGGCCAGGCTCAGAACAG 5324

Qy        1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSer 1853
                                                               ||||||   |||
Db        5325 ------------------------------------------CCTGGCCGACCTGGCAGT 5342

Qy        1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGln 1872
               |||   :::            ||||||   |||   |||||||||:::      ||||||
Db        5343 CATGGATATGTTCGCTCCCCTTCCCT---TCAGTAAGAACTCAGGAGACCATGTTGCAA 5399

Qy        1873 GlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluPro 1892
```

```
               |||||||||||||||         |||            :::||||||   ::::::|||
Db        5400 CAGAGACCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATCACACCTTTGGATCCA 5459

Qy        1893 SerLysProThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThr 1912
               :::            ||||||                     :::   |||
Db        5460 ACT------GCTCAGCTACGA---------------------ATCATGCCA--------- 5483

Qy        1913 PheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeu 1932
                                    |||   ||||||
Db        5484 ------------------CTGCCTGCTGGGGGC-------------------------- 5498

Qy        1933 MetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArg 1952

Db        5498 ------------------------------------------------------------ 5498

Qy        1953 AlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAla 1972
                                             |||            ||||||
Db        5499 ------------------------------CCTTCAATAAGCCAAGGCCTG--------- 5519

Qy        1973 SerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAla 1992
                                                         |||   |||   :::
Db        5520 ------------------------------------------CCAGCCTCCCGTTACAAC 5537

Qy        1993 ThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPro 2012
               |||   |||         |||      :::
Db        5538 ACTGCTGCGGATGCCCTGGCTGCTCTTGTG------------------------------ 5567

Qy        2013 AlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSer 2032
                        |||:::||||||   |||      :::   :::::::::|||      |||
Db        5568 ------GATGCTGCAGCTTCTGCACCCCAGATGGATGTGTCCAAAACAAAAGAGATATCT 5621

Qy        2033 IleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGly 2052
                           |||      |||                              :::
Db        5622 TCTCACAGGTATGAAACACCTAGC------------------------------GATGCT 5651

Qy        2053 ValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLys 2072
               :::|||   :::||||||   |||||||||:::         :::|||
Db        5652 ATTGAGGTGATAAGTCCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAGACCTAT 5711

Qy        2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
                     |||   :::      :::      |||   :::         :::      |||
Db        5712 CAGCCAGAGGTTGTTAAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATATGAAGGA 5771

Qy        2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSer 2112
               |||                           |||   |||   ||||||
Db        5772 CCA---------------------------TTACATCACTATCGACCA---------CAG 5795

Qy        2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGly--------------------- 2125
               |||   |||   ||||||      |||      |||
Db        5796 CAGGAATCACCATCTCCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGAATGGGG 5855

Qy        2126 ---ValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThr 2144
                  |||            |||::::::|||||||||   ||||||   ::::::||||||
Db        5856 CAAGTGCCCAGGACCCATCGGCTGATCACACTTGCTGATCACATCTGTCAAATTATCACA 5915

Qy        2145 GlnAspTyrThrArgHisHis---------------ProGlnGlnLeuSerAlaProLeu 2159
               ||||||:::   |||:::                  |||||||||         |||
Db        5916 CAAGATTTGCTAGAAATCAAGTTTCCTCGCAGACTCCCCAGCAG---------CCTCCT 5966

Qy        2160 ProAlaProLeuTyrSerPheProGlyAla------SerCysProValLeuAspLeuArg 2177
                  :::         :::   |||   |||      |||   ||||||         |||
Db        5967 ACTTCTACATTCCAGAACTCACCTTCTGCTTTGGTATCTACACCTGTG---------AGG 6017

Qy        2178 ArgProProSerAspLeuTyrLeuProProProAsp---------------HisGlyAla 2192
                        |||:::   |||   |||                        |||
Db        6018 ACTAAAACATCAAACCGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCATCAAAGA 6077

Qy        2193 Pro---AlaArgGlySerProHis---------------SerGluGlyGlyLysArgSer 2206
               |||   :::|||   ||||||                  |||   |||   :::
Db        6078 CCAGGTTCAAGGGTCTCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGGCCTGGA 6137

Qy        2207 ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSer 2226
                     |||   ::::::   |||               :::      ||||||:::|||
Db        6138 AAATCCCCAGAGAGGAGTCACGTC---------TCTTCGGAGCCCTACGAGCCCATCTCC 6188

Qy        2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyr 2246
               ||||||:::                  |||   :::               ||||||
Db        6189 CCACCCCAG------GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTCTTGTCT 6242

Qy        2247 ArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
               :::   |||   :::            |||      ::::::|||||||||:::   |||
Db        6243 CAGAGGGGCGCAGAGCCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGTATAAGC 6302

Qy        2267 GlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
                     |||:::||||||:::||||||   |||:::   |||   |||||||||||||||
Db        6303 TACTTGCCTTCATTCTTCACCAAGCTT---GAAAATACATCACCCATGGTTAAATCAAAG 6359

Qy        2287 LysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIle 2306
               ||||||||||||   :::|||||||||:::            :::   :::
Db        6360 AAGCAGGAGATTTTTCGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATGGCAGCT 6419

Qy        2307 SerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThr 2326
               :::||||||||||||||||||||||||:::||||||:::|||   :::|||   ::::::
Db        6420 GCTCAGCCAGGAACTGAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCAGTTAGC 6479

Qy        2327 TyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIle 2346
                  |||      :::      :::   ||||||   |||:::|||||||||   ||||||
Db        6480 TCTAGAGGCCATTCTTTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGACATTATC 6536

Qy        2347 ArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu------------------Ser 2360
               ||||||||||||||||||   :::|||         |||                  |||
Db        6537 AGGAAGGCTCTCATGGGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTCATGTCC 6596

Qy        2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380
                  |||:::                  |||      |||:::   |||:::
Db        6597 CAGCCTATGGGAGTA---------GTGCCTGGTACTGCCAACACCTCAGTT--------- 6638

Qy        2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
                     :::|||:::   :::   |||   :::            ||||||      |||
Db        6639 ------GTGACCAGTGGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCATTCAGGA 6692

Qy        2401 Gly------LysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAla 2418
               |||      |||   |||:::      :::   :::|||||||||:::|||||||||
Db        6693 GGAGTTTGCAAACCAAAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCTATA 6752

Qy        2419 Pro------GlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGlu 2436
               |||      |||            :::||||||   ||||||||||||||||||||||||
Db        6753 CCTGGGCAAGGCTACTTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCAGAA 6812

Qy        2437 GlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSer 2456
               ||||||   :::|||:::||||||               |||||||||||||||||||||
Db        6813 GGGGATTACCATAGGCAGACGCCA------GGGTGGGCCTGGGAAGACAGGCCCTCTTCA 6866

Qy        2457 AlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMet 2476
                  |||||||||   ||||||||||||||||||   ||||||:::         :::
Db        6867 ACAGGCTCAACTCAGTTTCCTTATAACCCTCTGACTATGCGGATG---------CTCAGC 6917

Qy        2477 AlaSerProProProProGlyLeuProAlaGlySerGly---ProLeuAlaGlyProHis 2495
               ::::::||||||   |||         |||   |||            |||   ||||||
Db        6918 AGTACTCCACCAACACCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCTCAC 6977

Qy        2496 ---------HisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluThr 2512
                              |||:::   ||||||   |||||||||   :::||||||||||||
Db        6978 CAACAGAACAGGATCTGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAGACC 7037

Qy        2513 LeuSerAspSerGlu 2517
               ||||||||||||:::
Db        7038 CTGTCGGATAGTGAT 7052

RESULT 11
MMU35312
LOCUS       MMU35312                7780 bp    mRNA    linear   ROD 19-OCT-1996
DEFINITION  Mus musculus nuclear receptor co-repressor mRNA, complete cds.
ACCESSION   U35312
VERSION     U35312.1  GI:1022717
KEYWORDS    .
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 7780)
  AUTHORS   Horlein,A.J., Naar,A.M., Heinzel,T., Torchia,J., Gloss,B.,
            Kurokawa,R., Ryan,A., Kamei,Y., Soderstrom,M., Glass,C.K. and
            Rosenfeld,M.G.
  TITLE     Ligand-independent repression by the thyroid hormone receptor
            mediated by a nuclear receptor co-repressor
  JOURNAL   Nature 377 (6548), 397-404 (1995)
  MEDLINE   96008539
   PUBMED   7566114
REFERENCE   2  (bases 1 to 7780)
  AUTHORS   Horlein,A.J., Neer,A.M., Heinzel,T., Torchia,J., Gloss,B.,
            Kurokawa,R., Kamai,Y., Ryan,A., Glass,C.K. and Rosenfeld,M.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1995) Andreas J. Horlein, School of Medicine,
            UCSD, 9500 Gilman Drive, La Jolla, CA 92093-0648, USA
FEATURES             Location/Qualifiers
     source          1. .7780
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /tissue_type="pituitary"
     CDS             117. .7478
                     /note="N-CoR"
                     /codon_start=1
                     /product="nuclear receptor co-repressor"
                     /protein_id="AAB17125.1"
                     /db_xref="GI:1022718"
                     /translation="MSSSGYPPNQGAFSTEQSRYPSHSVQYTFPSARHQQEFAVPDYR
                     SSHLEVSQASQLLQQQQQQLRRRPSLLSEFHPGSDRPQERRSGYEQFHPGPSPVDHD
                     SLESKRPRLEQVSDSHFQRISAAVLPLVHTLPEGLRSSANAKKDPAFGVKHEAPSSPL
                     SGQPCGDDQNASPSKLSKEELIQSMDRVDREIAKVEQQILKLKKKQQQLEEEAAKPPE
                     PEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKIFEGLGPKVELPLYNQPSDTKVYH
                     ENIKTNQVMRKKLILFFKRRNHARKQREQKICQRYDQLMEAWEKKVDRIENNPRRKAK
                     ESKTREYYEKQFPEIRKQREQQERFQRVGQRGAGLSATIARSEHEISEIIDGLSEQEN
                     NEKQMRQLSVIPPMMFDAEQRRVKFINMNGLMEDPMKVYKDRQFMNVWTDHEKEIFKD
                     KFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKALVRRNYGKRRGRNQQIARP
                     SQEEKVEEKEEDKAEKTEKKEEEKKDDEEKDDKEDSKETTKEKDRTEATAEEPEEREQ
                     VTPRGRKTANSQGRGKGRVTRSMTSEAAAANAAAAATEEPPPPLPPPPEPISTEPVET
                     SRWTEEEMEVAKKGLVEHGRNWAAIAKMVGTKSEAQCKNFYFNYKRRHNLDNLLQQHK
                     QKASRKPREERDVSQCESVASTVSAQEDEDIEASNEEENPEDSEGAENSSDTESAPSP
                     SPVEAAKSSEDSSENAASRGNTEPVAELEATTDPAPCASPSSAVPTTKPAERESVEAQ
                     VTDSASAETAEPMDVDHEECGAEGSSVLDPPAPTKADSVDPEMQVPENTASKGEGDAK
                     ERDLESTSEKTEARDEDVVVAEQIERPEPQSDDDSSATCSADEGVDGEPERQRVFPMD
                     AKPSLLTPPGSILISSPIKPNLLDLPQLQHRAAVIPPMVSCTPCNIPIGTPVSGYALY
                     QRHIKAMHESALLEEQRQRQEQVDLECRSSTSPCSTSKSPNREWEVLQPAPHQVITNL
                     PEGVRLPTTRPTRPPPPLIPSSKTTVASEKPSFIMGGSISQGTPGTYLSSHNQAYPQE
                     APKPSVGSISLGLPRQQESTKAAPLTYIKQEEFSPRSQNSQPEGLLVRAQHEGVVRGT
                     AGAVQEGSITRGTPASKISVETISSLRGSITQGTPALPQAGIPTEALVKGPVSRMPIE
                     ESSPEKVREEAASKGHVIYEGKSGHILSYDNIKNAREGTRSPRTAHEMSLKRSYEAVE
                     GSIKQGMSMRESPVSAPLEGLICRALPRGSPHSDLKERTVLSGSIMQGTPRATAESFE
                     DGLKYPKQIKRESPPIRAFEGAITKGKPYDGITTIKEMGRSIHEIPRQDILTQESRKT
                     PEVVQSTRPIIEGSISQGTPIKFDNNSGQSAIKHNVKSLITGPSKLPRGMLEIVPENI
                     KVVERGKYEDVKAGEPVRARHTSVVSSGPSVLRSTLHEAPKAQLSPGLYDDSSARRTP
                     VSYQNTISRGSPMMNRTSDVSSSKSASHERKSTLTPTQRESIPAKSPVPGVDPIVSHS
                     PFDPHHRSSAAGEVYRSHLPTHLDPAMPFHRALDPAAAYLLQRQLSPTPGYPSQYQLY
                     AMENTRQTILNDYITSQQMQVNLRPDVTRGLSPREQPLGLPYPATRGIIDLTNMPPTI
                     LVPHAGGTSTPPMDRITYIPGTQVTFPPRPYNAASLSPGHPTHLAAAASAERERERER
                     EKERERERERERERERIAAAPADLYLRPGSEQPGRPGSHGYVRSPSPSVRTQETIL
                     QQRPSVFQGTNGTSVITPLDPTAQLRIMPLPSGGPSISQGLPASRYNTAADALAALVD
                     AAASAPQMDVSKTKESKHEAARLEENLRSRSAAVSEQQQLEQKNLEVEKRSVQCVCTS
                     SALPSGKAQPHASVVYSEAGKDKGPPPKSRYEEELRTRGKTTITAANFIDVIITRQIA
                     SDKDARERGSQSSDSSSSLSSHRYETASDAIEVISPASSPAPPQEKPQAYQPDMVKAN
                     QAENESTRQYEGPLHHYRSQQESPSPQQQPPLPPSSQSEGMGQVPRTHRLITLADHIC
                     QIITQDFARNQVPSQASTSTFQTSPSALSSTPVRTKTSSRYSPESQSQTVLHPRPGPR
                     VSPENLVDKSRGSRPGKSPERSHIPSEPYEPISPPQGPAVHEKQDSMLLLSQRGVDPA
                     EQRSDSRSPGSISYLPSFFTKLESTSPMVKSKKQEIFRKLNSSGGGDSDMAAAQPGTE
                     IFNLPAVTTSGAVSSRSHSFADPASNLGLEDIIRKALMGSFDDKVEDHGVVMSHPVGI
                     MPGSASTSVVTSSEARRDEGEPSPHAGVCKPKLINKSNSRKSKSPIPGQSYLGTERPS
                     SVSSVHSEGDYHRQTPGWAWEDRPSSTGSTQFPYNPLTIRMLSSTPPTQIACAPSAIT
```

```
                     QAAPHQQNRIWEREPAPLLSAQYETLSDSDD"
ORIGIN

Alignment Scores:
Pred. No.:              1.68e-76        Length:         7780
Score:                  4187.00         Matches:        1104
Percent Similarity:     53.94%          Conservative:   347
Best Local Similarity:  41.04%          Mismatches:     798
Query Match:            31.68%          Indels:         442
DB:                     10              Gaps:           104

US-09-522-753-5 (1-2517) x MMU35312 (1-7780)

Qy         16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
              :::   |||||||||   ||||||:::   |||            ||||||
Db        165 CAGAGTCGCTACCCTTCGCACTCGGTTCAGTACACCTTTCCCAGCGCCCGTCACCAGCAG 224

Qy         36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
              :::      :::   :::|||:::      |||               ||||||   |||
Db        225 GAATTTGCAGTCCCTGACTACCGTTCTTCTCATCTTGAAGTCAGTCAGGCGTCACAGCTC 284

Qy         54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                |||      ||||||||||||||||||||||||||||||
Db        285 TTGCAGCAGCAGCAGCAGCAGCAGCTTCGCAGACGGCCTTCCCTGCTTTCCGAGTTTCAC 344

Qy         74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
              ||||||::::::|||   ||||||         |||               :::   |||
Db        345 CCGGGTTCCGACAGGCCCCAGGAA------AGGAGAAGTGGATACGAGCAGTTCCACCCG 398

Qy         94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                             :::   :::   :::||||||||||||||||||||||||   :::
Db        399 GGCCCTTCCCCGGTGGACCATGACTCGCTGGAGTCCAAGCGGCCTCGCCTGGAGCAGGTT 458

Qy        114 -ProAspProLeu------------LeuArgProSerProLeuLeuAlaThrGlyGlnPr 129
               |||   |||               :::   ||||||   |||      |||      ||
Db        459 TCCGACTCCCACTTCCAGCGCATCAGTGCTGCCGTCCTCCCTTTGGTGCACACG-CTGCC 517

Qy        129 oAlaGly---------SerGluAspLeuThrLysAspArgSerLeuThrGlyLysLeuGl 146
              |   |||         |||   :::      ||||||   :::         |||   ||
Db        518 AGAAGGACTGAGGTCTTCTGCCAATGCTAAGAAGGATCCGGCATTTGGAGTCAAACATGA 577

Qy        146 u---ProValSerProProSerProProHisThrAspProGluLeuGluLeuValProPr 165
              |   |||   ||||||   |||                  :::            |||
Db        578 AGCTCCTTCCTCTCCCCTCTCTGGGCAGCCATGCGGAGATGATCAGAATGCCTCACCTTC 637

Qy        165 oArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGluIleThrMe 185
               :::||||||||||||||||||||||||:::||||||||||||||||||||||||
Db        638 AAAACTGTCAAAGGAAGAGCTGATACAGAGCATGGATCGTGTCGATCGAGAAATTGCGAA 697

Qy        185 tValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGluGluGluAlaAl 205
               |||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db        698 AGTAGAACAGCAGATCCTTAAACTGAAAAAGAAGCAGCAACAGCTCGAAGAAGAAGCTGC 757

Qy        205 aLysProProGluProGluLysProValSerProProProIleGluSerLysHisArgSe 225
              ||||||||||||||||||||||||||||||||||||||||:::|||   |||||||||||
Db        758 TAAACCCCCAGAGCCTGAGAAGCCTGTGTCCCCTCCTCCCGTGGAGCAGAAGCACCGAAG 817

Qy        225 rLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAlaHisArgIleLe 245
              |:::|||||||||||||||||||||||||||||||||||||||   ||||||:::|||
Db        818 TATTGTCCAAATCATTTATGACGAGAATCGGAAAAAAGCAGAAGAAGCTCATAAAATATT 877

Qy        245 uGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGl 265
               |||||||||||||||:::||||||||||||||||||||||||||||||||||||:::
Db        878 TGAAGGTCTTGGCCCAAAAGTTGAACTGCCGCTCTACAACCAGCCGTCAGATACCAAGGT 937

Qy        265 nTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIleLeuTyrPheLy 285
               ||||||||||||||||||   ||||||   |||||||||||||||||||||:::|||||
Db        938 GTACCACGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATTTTATTTTTTAA 997

Qy        285 sArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAspGlnLe 305
              |||||||||||||||||||||||||   :::||||||   ||||||||||||||||||||
Db        998 AAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAACGTTATGATCAGCT 1057
```

```
Qy        305 uMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgArgAlaLy 325
              ||||||||||   ||||||||||||:::||||||||||||||||||||||||:::|||||
Db       1058 CATGGAAGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGGAGGAAAGCTAA 1117

Qy        325 sGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLysGlnArgGl 345
              ||||||||||   |||||||||||||||||||||||||||||||||||||||||||||||
Db       1118 AGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGAAAACAAAGAGA 1177

Qy        345 uLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeuSerMetSerAl 365
              |   |||||||||   |||   ||||||||||||||||||:::|||||||||   :::
Db       1178 ACAGCAAGAAAGATTTCAG---CGAGTTGGTCAGAGGGGAGCTGGTCTTTCAGCCACCAT 1234

Qy        365 aAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGluGlnGluAsnLe 385
               ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db       1235 TGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGTCTTTCTGAACAGGAGAATAA 1294

Qy        385 uGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAspAlaAspGlnGl 405
               |||||||||||||||||||||:::|||||||||||||||::::::||||||:::|||::
Db       1295 TGAGAAGCAAATGCGTCAGCTTTCTGTGATTCCACCTATGATGTTTGATGCAGAACAAAG 1354

Qy        405 nArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAs 425
              :|||:::|||||||||||||||||||||||||||   |||||||||||||||||||||||
Db       1355 AAGGGTCAAATTCATCAATATGAATGGGCTGATGGAGGATCCAATGAAGGTTTATAAAGA 1414

Qy        425 pArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArgGluLysPheMe 445
              |||||||   ||||||:::|||::::::   |||||||||   |||::::::||||||::
Db       1415 CAGACAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAGGACAAGTTTAT 1474

Qy        445 tGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGl 465
              :|||||||||||||||||||||||||||||||||:::||||||||||||:::|||   ::
Db       1475 CCAGCATCCAAAAAACTTTGGACTAATTGCATCCTATTTGGAAAGGAAGAGTGTTCCTGA 1534

Qy        465 uCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSerLeuValArgAr 485
              :|||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       1535 TTGTGTTTTATATTACTATTTAACCAAGAAAAATGAGAATTATAAGGCCCTCGTGAGAAG 1594

Qy        485 gSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGl 504
              |:::|||   :::|||||||||::::::||||||      :::      |||::::::::
Db       1595 GAATTATGGAAAACGCAGAGGCAGAAATCAGCAGATTGCCCGTCCCTCACAAGAAGAAAA 1654

Qy        504 nGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGl 524
              :   :::::::::                        :::|||:::|||   ||||||
Db       1655 AGTAGAAGAAAAG------------------------GAAGAGGATAAAGCAGAAAAAAC 1690

Qy        524 uLysGluLysGluAlaGluLys------GluGluGluLysProGluValGluAsnAspLy 542
               ::::::||||||   ||||||      :::|||||||||   :::   |||:::   ||
Db       1691 AGAGAAAAAGGAAGAAGAAAAGAAGGATGATGAAGAAAAAGATGATAAGGAAGACTCTAA 1750

Qy        542 sGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGl 562
              ||||         ::::::         :::   :::   |||:::   :::|||:::||
Db       1751 AGAAACAACCAAGGAAAAGGACAGAACGGAAGCCACAGCAGAAGAACCTGAAGAAAGAGA 1810

Qy        562 uAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyArgIl 582
              |   |||      :::||||||||||||||||||||||||||||||   |||||||||::
Db       1811 GCAGGTCACTCCAAGGGGGCGAAAGACTGCTAACAGCCAAGGCCGCGGAAGGGCCGGGT 1870

Qy        582 eThrArgSerMetAlaAsnGlu---------AlaAsnSerGluGluAlaIleThrProGl 599
              :||||||||||||   :::|||         ||||||:::      |||   |||   ::
Db       1871 CACCAGGTCGATGACAAGTGAAGCTGCAGCTGCCAATGCTGCTGCTGCAGCCACTGAGGA 1930

Qy        599 nGlnSerAlaGluLeu------------------AlaSerMetGluLeuAsnGluSerSe 613
              :            |||                     |||   |||      |||:::||
Db       1931 GCCCCCGCCACCCCTGCCGCCACCACCAGAGCCCATTTCTACAGAACCTGTTGAGACTTC 1990

Qy        613 rArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAs 633
              |||||||||||||||||||||||||   |||||||||||||||:::||||||||||||||
Db       1991 AAGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGCCTGGTAGAACATGGTCGTAA 2050

Qy        633 nTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLysAsnPheTy 653
              ||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||||||||||
Db       2051 CTGGGCAGCCATTGCTAAAATGGTGGGAACTAAAAGTGAAGCCCAGTGCAAAAACTTCTA 2110

Qy        653 rPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMe 673
              |||||||||||||:::|||   |||||||||   :::|||||||||||||||   |||
Db       2111 TTTTAACTATAAAAGACGGCATAATCTTGACAACCTTTTGCAGCAACATAAACAGAAAGC 2170

Qy        673 tGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAlaSerGluGl 693
                  ::::::            |||         :::      :::   ||||||
Db       2171 TTCACGGAAACCCCGTGAGGAACGAGATGTATCTCAGTGTGAAAGTGTTGCTTCCACTGT 2230

Qy        693 uAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyValSerGlyAs 713
               :::|||               |||||||||::::::|||||||||            ||
Db       2231 TTCTGCC---------------CAAGAGGATGAAGATATTGAAGCCTCA------------AA 2266

Qy        713 nGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProAr 733
              ||||||||||      |||::::::|||
Db       2267 TGAGGAGGAAAATCCAGAAGATAGTGAAGGT----------------------------- 2297

Qy        733 gGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSe 753
                                             |||||||||||||||||||||   |||||
Db       2298 -------------------------GCTGAAAATAGTTCTGATACAGAAAGTGCTCCCTC 2332

Qy        753 rProHis---ThrGluAlaAlaLys---------AspThrGlyGlnAsnGlyProLysPr 769
              ||||         ||||||||||||         |||:::   :::|||
Db       2333 TCCTTCACCAGTTGAAGCTGCCAAGTCCAGTGAAGACAGCAGTGAAAATGCT-------- 2384

Qy        769 oProAlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgTh 789

Db       2385 ----------------------------------------------------------GC 2386

Qy        789 rSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProPr 809
               ||||||         ||||||            |||||||||      |||   |||
Db       2387 TTCTCGAGGAAACACCGAGCCTGTGGCTGAGCTTGAGGCCACCACTGACCCTGCACCCTG 2446

Qy        809 oProAlaProProSerProSerAlaProProProValValProLysGluGluLysGluGl 829
                  :::|||         ||||||   |||         |||   |||   :::
Db       2447 TGCATCTCCC------TCTTCAGCAGTTCCAACCACAAAACCAGCAGAAAGGGAAAGCGT 2500

Qy        829 uGlu----------------------ThrAlaAlaAlaProProValGluGluGlyGl 841
               |||                                 |||   |||::::::
Db       2501 GGAGGCCCAGGTGACCGACAGCGCCAGTGCCGAGACCGCAGAGCCGATGGACGTAGACCA 2560

Qy        841 uGluGln---------------------------LysProProAlaAlaGluGluLeuAl 852
               |||:::                              |||||||||      :::
Db       2561 TGAGGAGTGCGGTGCCGAGGGCAGTTCTGTTCTTGATCCACCAGCCCCTACCAAA----- 2615

Qy        852 aValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAlaGluGl 872
                  |||:::         |||      |||      |||   |||            ||
Db       2616 -GCCGACTCCGTGGACCCAGAAATGCAGGTA---CCAGAAAATACTGCGTCTAAAGGTGA 2671

Qy        872 uGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLy 892
              ||||   ||||||   :::|||   |||:::
Db       2672 AGGGGATGCCAAGGAAAGAGACTTGGAGAGCACC-------------------------- 2705

Qy        892 sAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaPr 912
               :::||||||   |||                        |||:::            ||
Db       2706 -AGTGAGAAGACAGAGGCTAGAGATGAAGACGTGGTAGTGGCTGAACAGATCGAGAGGCC 2764

Qy        912 oGln---------AspSerAspSerSerAlaThrCysSerAlaAspGluValAspGluAl 929
              |:::         |||   ||||||||||||||||||||||||||||||      :::
Db       2765 TGAGCCACAGTCAGACGACGACTCCAGTGCCACTTGCAGTGCGGATGAGGGTGTGGATGG 2824

Qy        929 aGluGlyGlyAspLysAsnArgLeu------LeuSerProArgProSerLeuLeuThrPr 947
               |||      ::::::   |||:::      :::      :::|||||||||||||||||
Db       2825 AGAG---CCAGAGAGGCAGAGAGTGTTTCCCATGGATGCAAAGCCTTCATTGTTAACTCC 2881

Qy        947 oThrGlyAspProArgAlaAsnAlaSerProGlnLysPro------LeuAspLeuLysGl 965
              |   |||               :::||||||   ||||||      |||||||||   ||
Db       2882 TCCTGGATCTATCCTAATC---TCATCCCCTATTAAACCAAACCTATTGGATCTGCCACA 2938

Qy        965 nLeuLysGlnArgAlaAlaAlaIleProProIle------------------------ 976
              ||||:::   |||||||||   |||||||||:::
Db       2939 GCTTCAGCATCGAGCTGCTGTTATTCCACCAATGGTTTCTTGCACTCCATGTAATATACC 2998

Qy        977 ---------------------------------------GlnValThrLysValHisGl 983
                                                         :::      :::|||||
Db   2999 AATTGGAACGCCCGTAAGTGGCTATGCTCTTTACCAACGGCACATTAAGGCCATGCATGA 3058

Qy    983 uProProArgGluAspAlaAlaProThrLysProAlaProProAlaProProProProGl 1003
          |            :::                                          ::
Db   3059 GTCAGCACTCCTGGAG----------------------------------GAGCAGCG 3082

Qy   1003 nAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSe 1023
          :      |||   :::   |||         :::      :::||||||         ||
Db   3083 GCAGAGGCAAGAACAGGTAGACTTGGAATGCAGAAGCTCTACAAGCCCATGCAGCACTTC 3142

Qy   1023 rArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuPr 1043
          |:::||||||            ::::::|||
Db   3143 TAAGAGTCCA------------AACAGGGAG---------------------------- 3161

Qy   1043 oGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArgGluValIl 1063
                          |||                                 ||||||::
Db   3162 ----------------TGG---------------------------------GAAGTCCT 3172

Qy   1063 eLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisPr 1083
          ::::   :::||||||                     :::      |||   |||
Db   3173 CCAACCTGCTCCGCAT------------CAAGTGATAACTAACCTTCCTGAAGGGGTTCG 3220

Qy   1083 oLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProProThrIleSe 1103
           ||||||               |||   ||||||      |||   ||||||   |||
Db   3221 GCTTCCA---------------ACAACACGACCAACCAGGCCACCACCTCCCCTCATCCC 3265

Qy   1103 rAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGl 1123
           :::            ::::::|||   |||   ||||||   :::            ||
Db   3266 ATCATCTAAAACCACAGTGGCATCAGAAAAA---CCATCCTTTATA------ATGGGAGG 3316

Qy   1123 yAlaIleSerGlnGly---------------MetSerValGlnLeuHisValProTyrSe 1138
          |:::||||||||||||               :::|||
Db   3317 GTCTATCTCACAGGGAACTCCTGGCACTTACTTGTCTTCTCATAATCAGGCTTATCCACA 3376

Qy   1138 rGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPr 1158
           |||      |||      ||||||   ::::::::::|||||||||         :::
Db   3377 AGAAGCCCCTAAGCCCTCCGTGGGGTCTATCTCTCTGGGATTGCCCCGGCAGCAGGAGTC 3436

Qy   1158 oLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAl 1178
              |||   ||||||   :::   :::|||||||||:::   |||||||||   |||
Db   3437 TACCAAAGCAGCTCCTTTGACCTACATCAAGCAGGAAGAATTTTCTCCGAGAAGCCAAAA 3496

Qy   1178 aGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyTh 1198
                 ||||||   |||   |||      ||||||         |||:::||||||||
Db   3497 CTCACAACCTGAGGGTTTATTGGTC---AGAGCGCAGCATGAAGGTGTGGTCAGAGGCAC 3553

Qy   1198 rAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPr 1218
          ||||   |||:::|||      ||||||||||||:::|||   |||::::::::::::
Db   3554 TGCA---GGGGCTGTCCAAGAAGGAAGTATAACTCGGGGAACTCCAGCCAGCAAAATCTC 3610

Qy   1218 oSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAla-------- 1235
              ::::::      :::   |||||||||||||||   ||||||||||||
Db   3611 AGTGGAGACCATTTCATCGTTGCGGGGCTCTATTACCCAGGGGACCCCAGCTCTGCCCCA 3670

Qy   1236 ----------------AspValLeuTyrLysGlyThrIleThrArgIle---IleGlyGl 1249
                          :::   |||   ||||||   ::::::|||:::   |||   ||
Db   3671 GGCTGGAATACCAACAGAGGCTTTGGTGAAGGGACCTGTCTCCAGGATGCCTATTGAAGA 3730

Qy   1249 uAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIl 1269
          |   ||||||   ::::::         ||||||::::::      ||||||||||||||
Db   3731 AAGCAGTCCTGAGAAGGTC---------AGAGAGGAAGCTGCATCCAAGGGCCATGTTAT 3781

Qy   1269 eTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCy 1289
          |||||||||||||   ||||||:::|||||||||:::         :::
Db   3782 CTATGAAGGCAAAAGTGGACATATCTTATCATATGATAATATTAAGAATGCC-------- 3833

Qy   1289 sSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysAr 1309
                    :::|||      |||            ||||||   :::      |||||
Db   3834 -------CGAGAAGGGACTCGGAGTCCAAGAACAGCTCATGAAATGAGTTTA---AAAAG 3883

Qy   1309 gThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer----------------- 1323
          |:::|||:::   :::||||||   :::   :::   :::|||
Db   3884 AAGCTATGAGGCAGTGGAAGGAAGTATAAAGCAAGGCATGTCGATGAGGGAGTCTCCTGT 3943

Qy   1324 -SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerProHi 1343
           ||||||   :::|||||||||:::   ||||||:::|||      |||   ||||||||
Db   3944 GTCAGCACCTTTAGAGGGTCTGATATGCCGAGCATTACCC------AGGGGGAGCCCTCA 3997

Qy   1343 s---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSe 1362
          |      |||||||||:::      :::   |||||||||   ||||||   ||||||::
Db   3998 TTCTGACCTCAAAGAGAGGACTGTGCTGTCTGGTTCCATAATGCAGGGCACACCAAGAGC 4057

Qy   1362 rTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArgGluGl 1382
          :      |||:::   ||||||   |||   :::      |||   :::|||||||||
Db   4058 CACAGCGGAAAGCTTTGAGGACGGCCTT---AAATACCCCAAACAGATAAAAAGGGAGAG 4114

Qy   1382 yThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLe 1402
                          ||||||   |||            |||
Db   4115 C---------------CCTCCCATCCGAGCATTTGAAGGTGCCATT-------------- 4145

Qy   1402 uGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGl 1422
                    |||   ||||||   ::::::|||   :::   |||:::||||||   ||
Db   4146 -------ACCAAAGGAAAACCA---TACGATGGT---ATCACCACCATCAAAGAAATGGG 4192

Qy   1422 yArgSerIleHisGluIleProArgGluGluLeu---------------ArgHisThrPr 1437
          |||||||||||||||||||||||||:::::::::               |||   |||||
Db   4193 GCGCTCCATCCATGAAATCCCACGGCAAGATATTCTAACTCAGGAAAGCCGGAAAACTCC 4252

Qy   1437 oGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLe 1457
          ||||:::      :::   ||||||:::   ||||||||||||:::||||||||||||::
Db   4253 AGAAGTGGTCCAGAGCACGAGGCCAATAATTGAGGGTTCCATTTCCCAGGGCACACCAAT 4312

Qy   1457 uLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIl 1477
          :|||:::|||         |||         :::   ||||||:::|||:::||||||||
Db   4313 AAAATTTGACAACAAC---TCAGGTCAATCAGCTATCAAACACAATGTGAAGTCCTTAAT 4369

Qy   1477 eGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAsp---Al 1496
          |      |||   :::         |||         |||:::::::::   :::
Db   4370 CACAGGGCCTAGCAAACTA------CCCCGTGGAATGCTGGAAATTGTACCAGAGAACAT 4423

Qy   1496 aArgAlaLeuGluArgAlaCysTyrGlu---------------GluSerLeuLysSerAr 1511
           :::   :::||||||      ||||||               |||   :::::::::||
Db   4424 AAAAGTAGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAGCCAGTGCGAGCCCG 4483

Qy   1511 gProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValPr 1531
          |      :::      |||||||||   |||:::   |||               :::
Db   4484 GCACACGTCAGTGGTGAGCTCTGGCCCCTCCGTTCTCAGGTCTACA---------CTTCA 4534

Qy   1531 oGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAla-------- 1548
           |||      |||   :::   ||||||      |||:::|||      |||
Db   4535 CGAAGCTCCCAAAGCACAGCTGAGCCCGGGACTCTATGATGACAGCAGCGCTCGCAGGAC 4594

Qy   1549 -------ProPheAlaGlyHisLeuProArgGlySerProValThrMetArgGluProTh 1566
                    :::         :::   ||||||||||||:::      |||
Db   4595 TCCTGTGAGCTACCAGAACACCATATCCAGAGGCTCCCCCATGATGAACAGA-------- 4646

Qy   1566 rProArgLeuGlnGluGlySerLeuSerSerSerLys---AlaSerGlnAspArgLysLe 1585
                                :::||||||||||||   ||||||   :::||||||
Db   4647 -------------ACTTCTGATGTTTCTTCCAGCAAGTCTGCCAGTCATGAAAGGAAATC 4693

Qy   1585 uThrSerThrProArgGlu------------IleAlaLysSerProHisSerThrValPr 1601
           |||   ||||||   :::               ||||||||||||         |||
Db   4694 AACTCTGACCCCAACCCAAAGAGAAAGTATACCAGCCAAGTCTCCAGTGCCCGGGGTGGA 4753

Qy   1601 oGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyVa 1621
                       |||      ||||||::::::         |||      :::
Db   4754 TCCCATCGTGAGTCAC------AGCCCATTTGATCCTCATCACAGGAGTAGCGCTGCAGG 4807

Qy   1621 lAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGlyIl 1641
           ::::::||||||||||||:::|||         ||||||   ::::::|||
Db   4808 AGAGGTTTATCGGAGCCACCTACCCACGCACTTGGATCCA---GCTATGCCCTTTCACAG 4864

Qy   1641 eProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTy 1661
              ||||||   ||||||||||||   |||   |||   |||:::|||   |||   ||
Db   4865 GGCTTTGGATCCTGCTGCTGCTTACCTGTTACAGAGACAGCTTTCACCAACCCCAGGATA 4924
```

```
Qy    1661 rProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAs 1681
           ||||      |||      |||                           |||:::|||||
Db    4925 CCCAAGTCAGTACCAGCTCTAT---------------------------GCAATGGAGAA 4957

Qy    1681 n---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAl 1700
           |   ||||||||||||:::|||||||||||||||||||||||||||      |||
Db    4958 TACAAGGCAGACAATCCTCAACGATTACATTACCTCACAGCAGATGCAGGTGAATCTG-- 5015

Qy    1700 aThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLe 1720
                           |||   |||:::   |||||||||||||||||||||      ||
Db    5016 ----------------CGCCCTGATGTCACCAGGGGACTGTCCCCACGAGAGCAGCCACT 5059

Qy    1720 uAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLe 1740
           |   |||   |||   |||      ||||||||||||||||||:::   :::|||
Db    5060 GGGGCTCCCTTACCCAGCT---ACAAGAGGAATCATTGACCTGACCAATATGCCTCCA-- 5114

Qy    1740 uProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTy 1760
               :::|||||||||         ||||||            |||||||||:::   ||
Db    5115 -ACAATCTTAGTGCCTCATGCAGGGGGAACGAGCACCCCTCCCATGGACAGGATCACGTA 5173

Qy    1760 rLeuProThrAlaProGlnProPheSerSerArg---HisSerSerSerProLeuSerPr 1779
           |:::|||               |||      |||   ::::::::::::   ||||||||
Db    5174 TATTCCTGGTACACAGGTTACTTTCCCTCCCAGGCCATATAACGCTGCTTCTCTGTCTCC 5233

Qy    1779 oGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAs 1799
           ||||   ||||||||||||                  :::|||:::||||||||||||::
Db    5234 AGGACACCCAACACACCTT---------GCAGCAGCTGCAAGTGCTGAGAGGGAACGAGA 5284

Qy    1799 pArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrTh 1819
           :|||:::||||||::::::|||:::||||||||||||:::
Db    5285 ACGGGAAAGGGAGAAGGAGCGCGAACGTGAGCGCGAGCGTGAGCGTGAACGTGAACGCGA 5344

Qy    1819 rThrValGluHisAlaPro---------IleTrpArgProGlyThrGluGlnSerSerGl 1836
               :::      ||||||               |||||||||:::||||||
Db    5345 AAGGATCGCTGCTGCTCCCGCTGACCTCTACCTACGACCAGGTTCAGAACAG-------- 5396

Qy    1836 ySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHi 1856
                                                   ||||||   ||||||   ::
Db    5397 ----------------------------------CCAGGCCGCCCTGGCAGCCACGGATA 5422

Qy    1856 sAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGlnGlnArgPr 1875
           :            ||||||   |||   |||||||||:::      ||||||||||||||
Db    5423 TGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAGGAGACCATCCTGCAACAGAGACC 5479

Qy    1875 oSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysPr 1895
           |||||||         |||            :::||||||   ::::::|||:::
Db    5480 CAGTGTTTTCCAGGGCACCAATGGAACCAGTGTAATCACACCTTTGGACCCAACTGCTCA 5539

Qy    1895 oThrValLeuArgSerThrSerThrSerSerProVal------------ArgProAlaAl 1911
                  :::            :::      |||                  ||||||::
Db    5540 GCTACGCATCATGCCACTGCCTTCTGGGGGCCCTTCCATAAGTCAAGGCCTGCCAGCCTC 5599

Qy    1911 aThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProTh 1931
           :   :::      |||            |||         :::|||
Db    5600 CCGTTACAACACTGCTGCGGATGCC---CTGGCTGCTCTTGTGGATGCTGCAGCTTCTGC 5656

Qy    1931 rLeuMetGluProValLeuLeuProLysGluAlaProArg-------------------- 1944
                        |||         ||||||:::
Db    5657 ACCCCAGATGGATGTTTCCAAAACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTAGAAGA 5716

Qy    1945 -------------------ValAlaArgProGluArgProArgAlaAspThrGlyHisAl 1958
                                 |||      |||:::   :::   :::      :::
Db    5717 AAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAAACCT 5776

Qy    1958 aPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGl 1978
               :::   |||         :::               ::::::   ||||||
Db    5777 GGAGGTGGAGAAGAGATCTGTTCAGTGTGTGTGCACTTCTTCAGCCCTTCCAAGTGGCAA 5836

Qy    1978 ySerGluProArgProLeuVal---------------------------------ProPr 1987
            ::::::|||         |||                                 |||||
Db    5837 GGCCCAGCCTCATGCCTCAGTAGTGTATTCTGAGGCTGGGAAAGATAAAGGGCCTCCTCC 5896

Qy    1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn------LeuAlaProHi 2005
           |   |||   :::            ||||||      |||            |||   ::
Db    5897 AAAATCCAGATATGAGGAAGAGCTAAGGACCCGAGGGAAGACTACCATTACTGCAGCTAA 5956

Qy    2005 sHisAlaSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLy 2025
           :                              ||||||      |||
Db    5957 CTTCATAGACGTGATCATCACCCGGCAAATTGCCTCGGACAAGGATGCGAGGGAACGTGG 6016

Qy    2025 sThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGl 2045
            :::||||||         |||                     ||||||      |||
Db    6017 CTCTCAAAGTTCAGACTCTTCT------------------AGTAGCTTGTCTTCTCACAG 6058

Qy    2045 ySerSerTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuTh 2065
                     :::   :::   :::|||   :::||||||   |||||||||:::
Db    6059 GTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCCCGCCAGCTCACCTGCACCACC 6118

Qy    2065 rHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGl 2085
               :::|||                  :::   :::      :::      |||   ||
Db    6119 CCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAAGGCAAATCAAGCAGAAAATGA 6178

Qy    2085 uLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHi 2105
           |         :::      ||||||                           |||   ||
Db    6179 GTCCACTCGACAGTATGAAGGTCCA---------------------------CTGCATCA 6211

Qy    2105 sLeuArg---------------ProLeuProGluSerGlnProSerSerSerProLeuLe 2120
           |   |||               |||   |||:::   ||||||         |||
Db    6212 TTATCGGTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTGCCCCCATCTTC 6271

Qy    2120 uGlnThr------AlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHi 2138
            |||:::               |||            |||::::::|||||||||   ||
Db    6272 CCAGTCAGAGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACACTTGCTGACCA 6331

Qy    2138 sIleSerGluValIleThrGlnAspTyrThrArgHisHis---ProGlnGlnLeuSerAl 2157
           ||||   ::::::||||||||||||:::   |||:::      |||   |||   |||
Db    6332 CATCTGTCAAATTATCACACAAGATTTTGCTAGAAATCAAGTTCCCTCGCAGGCTTCTAC 6391

Qy    2157 aProLeu------------ProAlaProLeuTyrSerPheProGlyAlaSerCysProVa 2173
                              |||:::   |||   |||   |||
Db    6392 TTCTACATTCCAAACTTCACCATCTGCTTTGTCATCCACACCT----------------- 6434

Qy    2173 lLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAsp----------- 2189
                  :::|||         |||      |||   |||
Db    6435 -------GTAAGAACTAAAACCTCAAGCCGCTACAGCCCAGAATCACAGTCTCAGACTGT 6487

Qy    2190 ----HisGlyAlaProAla---ArgGlySerProHis------------------SerGluGl 2202
               |||      |||      |||   ||||||                  |||   ||
Db    6488 CTTGCATCCCAGACCAGGTCCTAGAGTCTCTCCAGAAAATCTTGTGGATAAATCCCGGGG 6547

Qy    2202 yGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIl 2222
           |   :::            |||   ::::::   :::               :::
Db    6548 AAGCAGGCCTGGAAAATCTCCAGAGAGGAGTCATATC---------CCATCAGAGCCCTA 6598

Qy    2222 eGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTy 2242
            ||||||:::|||||||||:::|||               |||   :::
Db    6599 TGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTG------CATGAGAAGCAGGACAGCAT 6652

Qy    2242 rProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerPr 2262
               ||||||   :::   |||                  |||      |||:::|||||
Db    6653 GTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAGTGATTCTCGATCACC 6712

Qy    2262 oGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMe 2282
           ||||:::   |||      |||:::||||||:::||||||   ||||||   |||   ||
Db    6713 AGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAGCACATCACCCAT 6769

Qy    2282 tValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluPr 2302
           |||||||||||||||||||||||||   :::|||||||||:::            :::
Db    6770 GGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGGTGGAGGTGACTC 6829

Qy    2302 oGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyTh 2322
            :::         :::||||||||||||||||||||||||:::||||||:::|||   ::
Db    6830 TGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGCAGTTACCACATC 6889

Qy    2322 rGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLe 2342
           :|||   ::::::   ||||||   :::      :::   ||||||   |||:::|||||
Db      6890 AGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGT---AACCTTGGTCT 6946

Qy      2342 uGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu-------- 2359
             ||||   ||||||||||||||||||||||||   :::|||         |||
Db      6947 AGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGAAGATCATGG 7006

Qy      2360 ----------SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSe 2376
                       |||   |||:::                  |||   :::||||||   ||
Db      7007 TGTTGTCATGTCCCATCCTGTG---------GGCATTATGCCTGGTAGTGCCAGCACCTC 7057

Qy      2376 rLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSe 2396
             |:::               :::|||:::::::::   |||   |||            ||
Db      7058 AGTG---------------GTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCATC 7102

Qy      2396 rProGlyGlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLy 2415
             ||||      |||      |||   |||:::      :::   :::|||||||||:::||
Db      7103 ACCTCATGCAGGAGTATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCTAA 7162

Qy      2415 sSerProAlaProGlyLeuAla------SerGlyAspArgProProSerValSerSerVa 2433
             |||||||   ||||||   :::            :::||||||   ||||||||||||||
Db      7163 ATCTCCTATTCCTGGGCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCTGT 7222

Qy      2433 lHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspAr 2453
             ||||||||||||||||   :::|||:::||||||               |||||||||||
Db      7223 GCATTCAGAAGGTGATTACCACAGGCAGACACCA------GGATGGGCATGGGAAGATCG 7276

Qy      2453 gProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnAl 2473
             ||||||||||   |||||||||   ||||||||||||||||||   :::|||:::
Db      7277 GCCCTCTTCAACAGGTTCTACTCAGTTCCCTTACAACCCTCTGACCATACGGATG----- 7331

Qy      2473 aGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu----- 2491
                    ::::::|||   ||||||      :::                  :::
Db      7332 -------CTCAGCAGTACACCACCTACACAGATCGCATGCGCCCCATCTGCCATCACCCA 7384

Qy      2492 -AlaGlyProHis---------HisAlaTrpAspGluGluProLysProLeuLeuCysSe 2508
              |||   ||||||               |||:::   ||||||   |||||||||   ::
Db      7385 AGCAGCTCCACATCAACAGAACCGCATCTGGGAGAGGGAGCCTGCCCCGCTCCTCTCAGC 7444

Qy      2508 rGlnTyrGluThrLeuSerAspSerGlu 2517
             :||||||||||||||||||||||||:::
Db      7445 GCAGTATGAGACACTGTCTGATAGTGAC 7472

RESULT 12
CQ847916
LOCUS       CQ847916                7940 bp    DNA     linear   PAT 19-AUG-2004
DEFINITION  Sequence 221 from Patent WO2004063362.
ACCESSION   CQ847916
VERSION     CQ847916.1  GI:51469512
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Glover,D., Bell,G., Frenz,L. and Midgley,C.
  TITLE     Cell cycle progression proteins
  JOURNAL   Patent: WO 2004063362-A 221 29-JUL-2004;
            Cyclacel Limited (GB)
FEATURES             Location/Qualifiers
     source          1..7940
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="AF044209"
ORIGIN

Alignment Scores:
Pred. No.:             1.07e-75      Length:        7940
Score:                 4147.50       Matches:       1099
Percent Similarity:    53.33%        Conservative:  341
Best Local Similarity: 40.70%        Mismatches:    785
Query Match:           31.38%        Indels:        475
DB:                    6             Gaps:          108
```

```
US-09-522-753-5 (1-2517) x CQ847916 (1-7940)

Qy        16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
             :::   ||||||||||||||||||:::   |||               |||
Db       289 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 348

Qy        36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
             :::      :::   :::|||:::      |||               ||||||   |||
Db       349 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 408

Qy        54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                               |||      ||||||||||||||||||||||||||||||
Db       409 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 468

Qy        74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
             ||||||::::::|||   ||||||         |||               :::   |||
Db       469 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 522

Qy        94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                            :::   :::   :::||||||||||||||||||||||||   :::
Db       523 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 582

Qy       114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                |||                                 ||||||:::   |||   |||
Db       583 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 642

Qy       123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                         |||      |||   |||   |||      ||||||   :::
Db       643 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 690

Qy       143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
             ||||||   |||   |||   ||||||   |||                  :::
Db       691 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 750

Qy       162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                   |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db       751 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 810

Qy       182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
             ||||||      |||||||||||||||   ||||||||||||||||||||||||||||||
Db       811 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 870

Qy       202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
             |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db       871 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 930

Qy       222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
             ||||||||||||:::|||||||||||||||||||||||||||||||||||||||   |||
Db       931 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAAAAAGCAGAAGAAGCT 990

Qy       242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
             |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db       991 CATAAAATTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 1050

Qy       262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIle 281
             ||||||:::   ||||||||||||||||||   ||||||   ||||||||||||||||||
Db      1051 GATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATT 1110

Qy       282 LeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArg 301
             |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db      1111 TTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAGCGT 1170

Qy       302 TyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArg 321
             |||||||||||||||||||||   ||||||||||||:::|||||||||||||||||||||
Db      1171 TATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGG 1230

Qy       322 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArg 341
             |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db      1231 AGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGA 1290

Qy       342 LysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeu 361
             ||||||||||||   |||||||||   |||      ||||||||||||||||||:::||||||
Db      1291 AAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGGCAGAGGGGAGCTGGTCTT 1347

Qy       362 SerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGlu 381
             |||   :::   ||||||||||||||||||:::|||||||||||||||||||||||||||
Db      1348 TCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGGCTCTCTGAG 1407

Qy       382 GlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAsp 401
             |||||||||   |||||||||||||||||||||:::|||||||||||||||::::::|||
Db      1408 CAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGATTCCACCTATGATGTTTGAT 1467

Qy       402 AlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLys 421
             |||:::|||:::|||:::|||||||||||||||||||||||||||   ||||||||||||
Db      1468 GCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCTTATGGAGGACCCTATGAAA 1527

Qy       422 ValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArg 441
             ||||||||||||||||||   ||||||:::|||::::::   |||||||||   |||:::
Db      1528 GTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAG 1587

Qy       442 GluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLys 461
             :::||||||:::|||||||||||||||||||||||||||||||||:::||||||||||||
Db      1588 GACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCATCATACTTGGAGAGGAAG 1647

Qy       462 ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSer 481
             :::|||   :::|||||||||||||||||||||||||||||||||||||||||||||:::
Db      1648 AGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAAAATGAGAATTATAAAGCC 1707

Qy       482 LeuValArgArgSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
             ||||||||||||:::|||   :::|||||||||::::::||||||      :::
Db      1708 CTCGTCAGAAGGAATTATGGGAAACGCAGAGGCAGAAACCAGCAAATTGCTCGACCCTCG 1767

Qy       501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
             |||:::::::::   :::::::::::::::         ::::::   :::   ||||||
Db      1768 CAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAAGCAGAAAAAACAGAAAAA 1821

Qy       521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                |||:::||||||:::   |||   ||||||:::|||::::::         ||||||
Db      1822 AAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAAGAAGACTCCAAAGAAAAT 1881

Qy       541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                ||||||         |||:::|||   |||   |||:::   |||:::   :::|||
Db      1882 ACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA---GAAGAAACTGAGGAA 1929

Qy       561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
             :::|||            :::|||||||||||||||||||||||||||||||||||||||
Db      1930 AGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGTCAGGGCCGCCGTAAGGGC 1989

Qy       581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
             ||||||||||||||||||   |||||||||   :::      |||
Db      1990 CGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGTGCTGCAGCCGCAGCGGCT 2049

Qy       601 Ser------------------------------AlaGluLeuAlaSerMetGluLeuAsn 610
             :::                                 |||      |||   |||
Db      2050 ACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCCATTTCTACAGAGCCTGTG 2109

Qy       611 GluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHis 630
             |||:::|||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db      2110 GAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGTCTAGTAGAACAT 2169

Qy       631 GlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLys 650
             ||||||||||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||
Db      2170 GGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAAAGTGAAGCTCAATGTAAA 2229

Qy       651 AsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLys 670
             |||||||||||||||||||||:::|||   |||||||||   :::|||||||||||||||
Db      2230 AACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAACCTCTTACAGCAGCATAAA 2289

Qy       671 LeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAla 690
                |||   ::::::               |||         :::      :::   |||
Db      2290 CAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCTCAATGTGAAAGTGTCGCT 2349

Qy       691 SerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyVal 710
             |||      :::|||               |||||||||::::::|||||||||
Db      2350 TCCACTGTTTCTGCT------------CAGGAGGATGAAGATATTGAAGCCTCC------ 2391

Qy       711 SerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGlu 730
                   ||||||||||||      |||::::::|||
Db      2392 ------AATGAAGAAGAAAATCCAGAAGACAGCGAA------------------------ 2421

Qy       731 ValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSer 750

Db      2421 ------------------------------------------------------------ 2421

Qy       751 IleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProPro 770
                               ||||||   |||
Db      2422 ---------------GTTGAAGCTGTCAAG------------------------------ 2436

Qy       771 AlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThrSer 790
                                        |||         :::|||         ||||||
Db      2437 ---------------------------CCCAGCGAGGACAGTCCTGAAAATGCTACTTCT 2469

Qy       791 ArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProPro 810
             |||         ||||||            |||            ||||||
Db      2470 CGAGGAAACACAGAACCTGCGGTTGAGCTTGAG------------CCCACCACGGAAACT 2517

Qy       811 AlaPro---ProSerProSer---AlaProProProValValProLysGluGluLysGlu 828
             ||||||      |||||||||   |||   |||         |||   |||::::::
Db      2518 GCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACAAAACCAGCTGAAGATGAAAGT 2577

Qy       829 GluGluThrAlaAlaAlaProProVal--------------------------------- 837
                ||||||               :::
Db      2578 GTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACAGCAGAGCAGATGGATGTAGAT 2637

Qy       838 ------------------GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeu 851
                               |||||||||            ||||||   |||   :::
Db      2638 CAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGATCCCCCACCCGCTACCAAA--- 2694

Qy       852 AlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAlaGlu 871
                   |||:::         |||      |||                  :::   |||
Db      2695 ---GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAAAACCATGCATCTAAAGTTGAA 2751

Qy       872 GluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeu 891
                         |||   :::|||   :::   |||
Db      2752 GGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC------------------------ 2787

Qy       892 LysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGlyAla 911
                :::||||||   |||                        |||:::         |||
Db      2788 ---AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTGGTAGCTCAGCAAATAAATGCC 2844

Qy       912 ------------ProGln---AspSerAspSerSerAlaThrCysSerAlaAspGluVal 926
                         ||||||   |||:::||||||||||||||||||||||||||||||
Db      2845 CAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCCACGTGCAGCGCTGATGAG--- 2901

Qy       927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro---------ArgProSer 943
                :::   :::|||      :::   :::   :::   |||         :::||||||
Db      2902 ---GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTTCCTATGGACTCAAAGCCTTCA 2958

Qy       944 LeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------ProLeu 961
             ||||||   |||||||||               :::||||||   |||      ||||||
Db      2959 CTGTTAAACCCCACTGGATCTATACTCGTC---TCATCTCCGTTAAAACCAAATCCACTG 3015

Qy       962 AspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIle--------------- 976
             ||||||   ||||||:::   |||||||||   |||||||||:::
Db      3016 GATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCACCAATGGTATCCTGCACCCCA 3075

Qy       977 ---------------------------------------------------GlnValThr 979
                                                                   :::
Db      3076 TGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCTCTCTACCAGCGACACATTAAA 3135

Qy       980 LysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaPro 999
                :::||||||            :::
Db      3136 GCAATGCATGAGTCAGCACTCCTGGAG--------------------------------- 3162

Qy      1000 ProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerPro 1019
                      :::      |||   :::   |||         :::      :::||||||
Db      3163 ---GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAATGTAGAAGTTCTACAAGTCCA 3219
```

```
Qy       1020 ArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAla 1039
                 |||   |||:::||||||            ::::::|||
Db       3220 TGTGGCACATCCAAGAGTCCA------------AACAGAGAG------------------ 3249

Qy       1040 GlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPro 1059
                                         |||
Db       3250 ---------------------------TGG------------------------------ 3252

Qy       1060 ArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPro 1079
                 ||||||::::::   :::||||||                     :::      |||
Db       3253 ---GAAGTCCTTCAGCCTGCTCCACAT------------CAATTGATAACTAATCTCCCT 3297

Qy       1080 ProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPro 1099
                 |||      ||||||               |||   ||||||      |||   |||
Db       3298 GAAGGCGTTCGGCTTCCG---------------ACAACTCGACCAACCAGGCCACCGCCC 3342

Qy       1100 ProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGlu 1119
              |||   |||   :::            ::::::|||   |||   ||||||   :::
Db       3343 CCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAAAAA---CCATCTTTTATA--- 3396

Qy       1120 ArgGlnIleGlyAlaIleSerGlnGly---------------MetSerValGlnLeuHis 1134
                       |||:::||||||||||||               ::::::
Db       3397 ---ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACTTATTTGACTTCTCATAATCAG 3453

Qy       1135 ValProTyrSerGluHisAlaLysAlaPro---ValGlyProValThrMetGlyLeuPro 1153
                    |||::::::            |||   ||||||   :::::::::|||||||||
Db       3454 GCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTAGGATCTATCTCTCTTGGACTGCCA 3513

Qy       1154 LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSer 1173
                       :::      |||   |||            :::|||||||||:::   |||
Db       3514 CGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTACATCAAGCAGGAAGAATTTTCT 3573

Qy       1174 ProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSer 1193
              ||||||   |||         ||||||   |||   |||      ||||||
Db       3574 CCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC---AGGGCCCAACATGAAGGT 3630

Qy       1194 ValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIlePro 1213
              |||:::||||||||||||   |||::::::      ||||||||||||:::|||   |||
Db       3631 GTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGAAGTATAACTCGGGGAACTCCA 3687

Qy       1214 SerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThr 1233
              ::::::::::::      :::|||      :::   |||||||||||||||   ||||||
Db       3688 ACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGGGGCTCTATCACTCAGGGCACC 3747

Qy       1234 ProAla------------------------AspValLeuTyrLysGlyThrIleThrArg 1245
              ||||||                        :::   |||   ||||||:::|||:::|||
Db       3748 CCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTGGTGAAGGGGTCCATTTCGAGA 3807

Qy       1246 IleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLys 1265
              :::      |||||||||            ::::::|||||||||::::::      |||
Db       3808 ATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGCAGAGAGGAAGCTGCATCCAAA 3861

Qy       1266 GlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSer 1285
              ||||||||||||||||||||||||   ||||||:::|||||||||:::         :::
Db       3862 GGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTGTCATATGATAATATTAAGAAT 3921

Qy       1286 ValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAla 1305
                                   :::|||      |||            ||||||   :::
Db       3922 GCC---------------CGAGAAGGGACTAGGAGTCCAAGAACAGCTCATGAAATCAGT 3966

Qy       1306 AlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer------ 1323
                    ||||||:::|||:::   :::||||||   :::   :::   :::|||
Db       3967 TTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATAAAGCAAGGGATGTCAATGAGG 4023

Qy       1324 ------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArg 1339
                          ||||||   :::|||||||||:::   ||||||:::|||      |||
Db       4024 GAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGCCGAGCATTACCC------AGG 4077

Qy       1340 HisSerProHis---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358
                 |||||||||      |||||||||:::      :::   |||||||||   ||||||
Db       4078 GGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTGTCTGGCTCCATAATGCAGGGG 4137

Qy       1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378
                 ||||||:::      |||:::   ||||||   |||   :::      |||   :::
Db       4138 ACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT---AAATATCCCAAACAAATT 4194

Qy       1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
              |||||||||                  ||||||   |||            |||
Db       4195 AAAAGGGAAAGT---------------CCTCCCATACGAGCATTTGAAGGTGCCATT--- 4236

Qy       1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
                                   |||   ||||||   ::::::|||   :::   |||:::
Db       4237 ------------------ACCAAAGGAAAACCA---TATGATGGC---ATCACCACCATC 4272

Qy       1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu--------------- 1433
              ||||||   |||||||||||||||||||||||||||:::::::::
Db       4273 AAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAAGATATTTTAACTCAGGAAAGT 4332

Qy       1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
              |||   |||||||||:::      :::   ||||||:::   ||||||||||||:::|||
Db       4333 CGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATAATTGAGGGTTCCATTTCCCAG 4392

Qy       1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
              |||||||||:::|||:::|||         |||         :::   ||||||:::|||
Db       4393 GGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAATCTGCCATCAAACACAATGTC 4449

Qy       1474 ArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
              :::|||||||||      |||   :::            :::   ||||||:::::::::
Db       4450 AAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGAATGCCTCCGCTGGAAATTGTG 4509

Qy       1494 AlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu---------------GluSer 1507
                 :::      :::   :::||||||      ||||||               |||:::
Db       4510 CCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAGACC 4569

Qy       1508 LeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaPro 1527
              ::::::||||||      :::      |||||||||   |||:::   |||
Db       4570 GTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCCTCCGTTCTTAGGTCCACA--- 4626

Qy       1528 ValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGly 1547
                    :::   |||      |||   :::   ||||||      |||:::|||
Db       4627 ------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCTGGGATTTATGATGACACCAGT 4680

Qy       1548 Ala---------------ProPheAlaGlyHisLeuProArgGlySerProValThrMet 1562
              |||                  :::         :::   ||||||||||||:::
Db       4681 GCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCCAGAGGCTCACCCATGATGAAC 4740

Qy       1563 ArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAsp 1582
              |||               :::                  :::         :::   :::
Db       4741 AGAACTTCTGATGTTACAATTCCT------------CCTAACAAGTCTACCAATCATGAA 4788

Qy       1583 ArgLysLeuThrSerThrProArgGlu------------IleAlaLysSerProHisSer 1598
              ||||||   |||   ||||||   :::               ||||||||||||
Db       4789 AGGAAATCGACACTGACCCCTACCCAGAGGGAAAGTATCCCAGCGAAGTCTCCAGTGCCT 4848

Qy       1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
                 |||               |||      ||||||::::::         ||||||
Db       4849 GGGGTGGACCCTGTCGTGAGCCAC------AGTCCGTTTGATCCCCATCACAGAGGCAGC 4902

Qy       1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
              :::      ::::::|||   ||||||:::|||         ||||||   ::::::|||
Db       4903 ACTGCAGGCGAGGTTTATTGGAGCCACCTGCCCACGCAATTGGATCCA---GCCATGCCT 4959

Qy       1639 ArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaPro 1657
                          ||||||   |||||||||||||||         |||   |||:::|||
Db       4960 TTTCACAGGGCTTTGGATCCTGCAGCGGCTGCTTACCTGTTTCAGAGACAGCTTTCACCA 5019

Qy       1658 AsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAla 1677
                 |||   ||||||      |||      |||
Db       5020 ACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC--------------------------- 5052

Qy       1678 AlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHis 1696
              |||:::||||||   ||||||||||||:::|||||||||||||||||||||||||||
Db       5053 GCAATGGAGAACACAAGACAGACAATCTTAAATGATTACATTACCTCACAACAGATGCAA 5112

Qy       1697 HisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArg 1716
                 |||                     |||   |||:::   ||||||||||||||||||
Db      5113 GTGAACTTG------------------CGTCCAGATGTGGCCAGAGGACTCTCCCCAAGA 5154

Qy      1717 GluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGln 1736
             |||      |||   |||   |||   |||      ||||||||||||||||||:::
Db      5155 GAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGAGGAATCATTGACCTGACCAAT 5211

Qy      1737 ValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAsp 1756
             :::|||         :::|||||||||         ||||||            ||||||
Db      5212 ATGCCTCCA---ACAATTTTAGTGCCTCATCCAGGGGGAACAAGCACTCCTCCCATGGAC 5268

Qy      1757 ArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg---HisSerSerSer 1775
             |||:::   |||:::|||               |||      |||   ::::::|||:::
Db      5269 AGAATCACTTATATTCCTGGTACACAGATTACTTTCCCTCCCAGGCCGTACAACTCTGCT 5328

Qy      1776 ProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGlu 1795
                :::|||||||||   ||||||||||||                  :::|||:::|||
Db      5329 TCCATGTCTCCAGGACACCCAACACACCTT---------GCAGCTGCTGCAAGTGCTGAG 5379

Qy      1796 ArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeu 1815
             |||||||||:::|||:::||||||::::::|||:::|||||||||         |||
Db      5380 AGGGAACGGGAACGGGAGCGGGAGAAGGAGCGGGAGCGGGAACGG---------ATTGCT 5430

Qy      1816 ThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSer 1835
                :::::::::   :::                  |||||||||:::||||||
Db      5431 GCAGCTTCCTCCGACCTC------------TACCTGCGGCCAGGCTCAGAACAG------ 5472

Qy      1836 GlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSer 1855
                                                       ||||||   ||||||
Db      5473 ------------------------------------CCTGGCCGACCTGGCAGTCATGGA 5496

Qy      1856 HisAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGlnGlnArg 1874
             :::            ||||||   |||   |||||||||:::      ||||||||||||
Db      5497 TATGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAGGAGACCATGTTGCAACAGAGA 5553

Qy      1875 ProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLys 1894
             |||||||||         |||            :::||||||   ::::::|||:::
Db      5554 CCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATCACACCTTTGGATCCAACTGCT 5613

Qy      1895 ProThrValLeuArgSerThrSerThrSerSerProVal------------ArgProAla 1910
                      :::                     |||                  ||||||
Db      5614 CAGCTACGAATCATGCCACTGCCTGCTGGGGGCCCTTCAATAAGCCAAGGCCTGCCAGCC 5673

Qy      1911 AlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPro 1930
             :::   :::      |||            |||         :::|||
Db      5674 TCCCGTTACAACACTGCTGCGGATGCC---CTGGCTGCTCTTGTGGATGCTGCAGCTTCT 5730

Qy      1931 ThrLeuMetGluProValLeuLeuProLys------------GluAlaProArgValAla 1946
                            |||         |||            ||||||   |||:::
Db      5731 GCACCCCAGATGGATGTGTCCAAAACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTAGAA 5790

Qy      1947 ArgProGluArgProArgAlaAsp------------------------------------ 1954
                      |||   |||:::
Db      5791 GAAAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAACC 5850

Qy      1955 ------------------------------------ThrGlyHisAlaPheLeuAlaLys 1962
                                                 |||      ||||||   :::
Db      5851 CTGGAGGTGGAGAAGAGATCTGTTCAGTGTTTATACACTTCTTCAGCCTTTCCAAGTGGC 5910

Qy      1963 ProProAlaArgSerGlyLeuGluProAlaSerSer---------ProSerLysGlySer 1979
                |||               :::|||   ||||||                  |||
Db      5911 AAGCCC---------------CAGCCTCATTCTTCAGTAGTTTATTCTGAGGCTGGGAAA 5955

Qy      1980 GluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAla 1999
             :::      |||      ||||||   |||   :::            ||||||
Db      5956 GATAAAGGGCCT------CCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCAGAGGG 6009

Qy      2000 LysAsn------LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSer 2017
             |||            |||   :::                              ||||||
Db      6010 AAGACTACCATTACTGCAGCTAACTTCATAGACGTGATCATCACCCGGCAAATTGCCTCG 6069

Qy      2018 AlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGlu 2037
                   |||               :::||||||         |||
Db      6070 GACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT--------------- 6114

Qy      2038 LeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSer 2057
                   ||||||      |||                  :::   :::|||   :::|||
Db      6115 ---AGTAGCTTATCTTCTCACAGGTATGAAACACCTAGCGATGCTATTGAGGTGATAAGT 6171

Qy      2058 ProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeu 2077
             |||   |||||||||:::         :::|||                  |||   :::
Db      6172 CCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAGACCTATCAGCCAGAGGTTGTT 6231

Qy      2078 AspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGly 2097
                   :::      |||   :::         :::      ||||||
Db      6232 AAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATATGAAGGACCA------------ 6279

Qy      2098 GlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSer 2117
                            |||   |||   ||||||            |||   |||   |||
Db      6280 ---------------TTACATCACTATCGACCA---------CAGCAGGAATCACCATCT 6315

Qy      2118 ProLeuLeuGlnThrAlaProGly------------------------ValLysGlyHis 2129
             |||      |||      |||                           |||
Db      6316 CCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGAATGGGGCAAGTGCCCAGGACC 6375

Qy      2130 GlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArg 2149
                |||::::::|||||||||   ||||||   ::::::||||||||||||:::   |||
Db      6376 CATCGGCTGATCACACTTGCTGATCACATCTGTCAAATTATCACACAAGATTTTGCTAGA 6435

Qy      2150 HisHis---------------ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyr 2164
             :::                  |||||||||         |||      :::
Db      6436 AATCAAGTTTCCTCGCAGACTCCCCAGCAG---------CCTCCTACTTCTACATTCCAG 6486

Qy      2165 SerPheProGlyAla------SerCysProValLeuAspLeuArgArgProProSerAsp 2182
             :::   |||   |||      |||   ||||||         |||         |||:::
Db      6487 AACTCACCTTCTGCTTTGGTATCTACACCTGTG---------AGGACTAAAACATCAAAC 6537

Qy      2183 LeuTyrLeuProProProAsp---------------HisGlyAlaPro---AlaArgGly 2196
                |||   |||                        |||      |||   :::|||
Db      6538 CGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCATCAAAGACCAGGTTCAAGGGTC 6597

Qy      2197 SerProHis---------------SerGluGlyGlyLysArgSerProGluProAsnLys 2211
             ||||||                  |||   |||   :::            |||   :::
Db      6598 TCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGGCCTGGAAAATCCCCAGAGAGG 6657

Qy      2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
             :::   |||               :::      ||||||:::|||||||||:::
Db      6658 AGTCACGTC---------TCTTCCGAGCCCTACGAGCCCATCTCCCCACCCCAG------ 6702

Qy      2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
                         |||   :::               ||||||   :::   |||   :::
Db      6703 GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTCTTGTCTCAGAGGGGCGCAGAG 6762

Qy      2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
                         |||      ::::::|||||||||:::   |||      |||:::|||
Db      6763 CCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGTATAAGCTACTTGCCTTCATTC 6822

Qy      2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
             |||:::||||||   |||:::   |||   |||||||||||||||||||||||||||
Db      6823 TTCACCAAGCTT---GAAAATACATCACCCATGGTTAAATCAAAGAAGCAGGAGATTTTT 6879

Qy      2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
             :::|||||||||:::            :::   :::         :::||||||||||||
Db      6880 CGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGAACT 6939

Qy      2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
             ||||||||||||:::||||||:::|||   :::|||   ::::::   |||      :::
Db      6940 GAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCAGTTAGCTCTAGAGGCCATTCT 6999

Qy      2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
                   :::   ||||||   |||:::|||||||||   |||||||||||||||||||||
Db      7000 TTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGACATTATCAGGAAGGCTCTCATG 7056

Qy      2352 GlyLysTyrAspGlnTrpGluGlu------------------SerProProLeuSerAla 2365
             |||   :::|||         |||                  |||   |||:::
Db      7057 GGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTCATGTCCCAGCCTATGGGAGTA 7116
```

```
Qy       2366 AsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAla 2385
                          |||      |||:::   |||:::               :::|||:::
Db       7117 ---------GTGCCTGGTACTGCCAACACCTCAGTT---------------GTGACCAGT 7152

Qy       2386 AlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGly------LysAla 2403
                 :::   |||   :::            ||||||      ||||||      |||
Db       7153 GGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCATTCAGGAGGAGTTTGCAAACCA 7212

Qy       2404 LysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro------GlyLeu 2421
              |||:::      :::   :::|||||||||:::|||||||||   |||      |||
Db       7213 AAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCTATACCTGGGCAAGGCTAC 7272

Qy       2422 AlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArg 2441
                       :::||||||   ||||||||||||||||||||||||||||||   :::|||
Db       7273 TTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCAGAAGGGGATTACCATAGG 7332

Qy       2442 ArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrPro 2461
              :::||||||               |||||||||||||||||||||   |||||||||
Db       7333 CAGACGCCA------GGGTGGGCCTGGGAAGACAGGCCCTCTTCAACAGGCTCAACTCAG 7386

Qy       2462 PheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProPro 2481
              ||||||||||||||||||   ||||||:::         :::   ::::::||||||
Db       7387 TTTCCTTATAACCCTCTGACTATGCGGATG---------CTCAGCAGTACTCCACCAACA 7437

Qy       2482 ProGlyLeuProAlaGlySerGly---ProLeuAlaGlyProHis---------HisAla 2497
              |||         |||   |||            |||   ||||||
Db       7438 CCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCTCACCAACAGAACAGGATC 7497

Qy       2498 TrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||:::   ||||||   |||||||||   :::||||||||||||||||||||||||:::
Db       7498 TGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAGACCCTGTCGGATAGTGAT 7557

RESULT 13
AX578069
LOCUS       AX578069                7940 bp    DNA     linear   PAT 08-JAN-2003
DEFINITION  Sequence 191 from Patent WO02081745.
ACCESSION   AX578069
VERSION     AX578069.1  GI:27647277
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J.,
            Connolly,T., Jackson,A., Bushnell,S.E. and Rawadi,G.
  TITLE     Genes involved in osteogenesis, and methods of use
  JOURNAL   Patent: WO 02081745-A 191 17-OCT-2002;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .7940
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     /note="Homo sapiens nuclear receptor co-repressor 1
                     (NCOR1), mRNA"
ORIGIN

Alignment Scores:
Pred. No.:              1.07e-75       Length:          7940
Score:                  4147.50        Matches:         1099
Percent Similarity:     53.33%         Conservative:    341
Best Local Similarity:  40.70%         Mismatches:      785
Query Match:            31.38%         Indels:          475
DB:                     6              Gaps:            108

US-09-522-753-5 (1-2517) x AX578069 (1-7940)

Qy         16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
              :::   ||||||||||||||||||:::   |||               |||
Db        289 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 348

Qy         36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
              :::      :::   :::|||:::      |||               ||||||   |||
Db        349 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 408

Qy         54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                |||      ||||||||||||||||||||||||||||||
Db        409 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 468

Qy         74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
              ||||||::::::|||   ||||||         |||               :::   |||
Db        469 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 522

Qy         94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                             :::   :::   :::||||||||||||||||||||||||   :::
Db        523 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 582

Qy        114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                 |||                                 ||||||:::   |||   |||
Db        583 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 642

Qy        123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                          |||      |||   |||   |||      ||||||   :::
Db        643 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 690

Qy        143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
              ||||||   |||   |||   ||||||   |||                  :::
Db        691 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 750

Qy        162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                    |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db        751 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 810

Qy        182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
              ||||||      |||||||||||||||   ||||||||||||||||||||||||||||||
Db        811 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 870

Qy        202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        871 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 930

Qy        222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||   |||
Db        931 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAAAAAGCAGAAGAAGCT 990

Qy        242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
              |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db        991 CATAAAATTTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 1050

Qy        262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIle 281
              ||||||:::   ||||||||||||||||||   ||||||   ||||||||||||||||||
Db       1051 GATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATT 1110

Qy        282 LeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArg 301
              |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db       1111 TTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAGCGT 1170

Qy        302 TyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArg 321
              |||||||||||||||||||||   ||||||||||||:::|||||||||||||||||||||
Db       1171 TATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGG 1230

Qy        322 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArg 341
              |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       1231 AGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGA 1290

Qy        342 LysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeu 361
              ||||||||||||   |||||||||   |||   ||||||||||||||||||:::||||||
Db       1291 AAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGGCAGAGGGGAGCTGGTCTT 1347

Qy        362 SerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGlu 381
              |||   :::   ||||||||||||||||||:::|||||||||||||||||||||||||||
Db       1348 TCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGGCTCTCTGAG 1407

Qy        382 GlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAsp 401
              |||||||||   |||||||||||||||||||||:::|||||||||||||||::::::|||
Db       1408 CAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGATTCCACCTATGATGTTTGAT 1467

Qy        402 AlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLys 421
              |||:::|||:::|||:::|||||||||||||||||||||||||||   ||||||||||||
Db       1468 GCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGCTTATGGAGGACCCTATGAAA 1527

Qy        422 ValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArg 441
              ||||||||||||||||||   ||||||:::|||::::::   |||||||||   |||:::
Db       1528 GTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAG 1587

Qy        442 GluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLys 461
              :::||||||:::|||||||||||||||||||||||||||||||||:::||||||||||||
Db       1588 GACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCATCATACTTGGAGAGGAAG 1647

Qy        462 ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSer 481
              :::|||   :::|||||||||||||||||||||||||||||||||||||||||||||:::
Db       1648 AGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAAAATGAGAATTATAAAGCC 1707

Qy        482 LeuValArgArgSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              ||||||||||||:::|||   :::|||||||||::::::||||||      :::
Db       1708 CTCGTCAGAAGGAATTATGGAAACGCAGAGGCAGAAACCAGCAAATTGCTCGACCCTCG 1767

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              |||:::::::::   :::::::::::::::         ::::::   :::   ||||||
Db       1768 CAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAAGCAGAAAAAACAGAAAAA 1821

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                 |||:::||||||:::   |||   ||||||:::|||::::::         ||||||
Db       1822 AAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAAGAAGACTCCAAAGAAAAT 1881

Qy        541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                 ||||||         |||:::|||   |||   |||:::   |||:::   :::|||
Db       1882 ACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA---GAAGAAACTGAGGAA 1929

Qy        561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
              :::|||            :::|||||||||||||||||||||||||||||||||||||||
Db       1930 AGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGTCAGGGCCGCCGTAAGGGC 1989

Qy        581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
              ||||||||||||||||||   |||||||||   :::      |||
Db       1990 CGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGTGCTGCAGCCGCAGCGGCT 2049

Qy        601 Ser------------------------------AlaGluLeuAlaSerMetGluLeuAsn 610
              :::                                 |||      |||   |||
Db       2050 ACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCCATTTCTACAGAGCCTGTG 2109

Qy        611 GluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHis 630
              |||:::|||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db       2110 GAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGTCTAGTAGAACAT 2169

Qy        631 GlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLys 650
              ||||||||||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||
Db       2170 GGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAAAGTGAAGCTCAATGTAAA 2229

Qy        651 AsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLys 670
              |||||||||||||||||||||:::|||   |||||||||   :::|||||||||||||||
Db       2230 AACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAACCTCTTACAGCAGCATAAA 2289

Qy        671 LeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAla 690
                 |||      ::::::            |||         :::      :::   |||
Db       2290 CAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCTCAATGTGAAAGTGTCGCT 2349

Qy        691 SerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyVal 710
              |||      :::|||               |||||||||::::::|||||||||
Db       2350 TCCACTGTTTCTGCT------------CAGGAGGATGAAGATATTGAAGCCTCC------ 2391

Qy        711 SerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGlu 730
                    ||||||||||||      |||::::::|||
Db       2392 ------AATGAAGAAGAAAATCCAGAAGACAGCGAA------------------------ 2421

Qy        731 ValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSer 750

Db       2421 ------------------------------------------------------------ 2421

Qy        751 IleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProPro 770
                                ||||||   |||
Db       2422 ---------------GTTGAAGCTGTCAAG------------------------------ 2436

Qy        771 AlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThrSer 790
                                         |||         :::|||         ||||||
Db       2437 ---------------------------CCCAGCGAGGACAGTCCTGAAAATGCTACTTCT 2469

Qy        791 ArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProPro 810
              |||         ||||||            |||            ||||||
Db       2470 CGAGGAAACACAGAACCTGCGGTTGAGCTTGAG------------CCCACCACGGAAACT 2517

Qy        811 AlaPro---ProSerProSer---AlaProProProValValProLysGluGluLysGlu 828
              ||||||      |||||||||   |||   |||         |||   |||::::::
Db       2518 GCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACAAAACCAGCTGAAGATGAAAGT 2577

Qy        829 GluGluThrAlaAlaAlaProProVal--------------------------------- 837
                 ||||||               :::
Db       2578 GTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACAGCAGAGCAGATGGATGTAGAT 2637

Qy        838 ------------------GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeu 851
                                |||||||||            ||||||   |||   :::
Db       2638 CAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGATCCCCCACCCGCTACCAAA--- 2694

Qy        852 AlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAlaGlu 871
                    |||:::         |||      |||                  :::   |||
Db       2695 ---GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAAAACCATGCATCTAAAGTTGAA 2751

Qy        872 GluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeu 891
                          |||   :::|||   :::   |||
Db       2752 GGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC------------------------ 2787

Qy        892 LysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGlyAla 911
                 :::||||||   |||                        |||:::         |||
Db       2788 ---AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTGGTAGCTCAGCAAATAAATGCC 2844

Qy        912 ------------ProGln---AspSerAspSerSerAlaThrCysSerAlaAspGluVal 926
                          ||||||   |||:::||||||||||||||||||||||||||||||
Db       2845 CAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCCACGTGCAGCGCTGATGAG--- 2901

Qy        927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro---------ArgProSer 943
                 :::   :::|||      :::   :::   :::   |||         :::||||||
Db       2902 ---GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTTCCTATGGACTCAAAGCCTTCA 2958

Qy        944 LeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------ProLeu 961
              ||||||   |||||||||               :::||||||   |||      ||||||
Db       2959 CTGTTAAACCCCACTGGATCTATACTCGTC---TCATCTCCGTTAAAACCAAATCCACTG 3015

Qy        962 AspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIle--------------- 976
              ||||||   ||||||:::   |||||||||   |||||||||:::
Db       3016 GATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCACCAATGGTATCCTGCACCCCA 3075

Qy        977 ---------------------------------------------------GlnValThr 979
                                                                    :::
Db       3076 TGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCTCTCTACCAGCGACACATTAAA 3135

Qy        980 LysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaPro 999
                 :::||||||            :::
Db       3136 GCAATGCATGAGTCAGCACTCCTGGAG--------------------------------- 3162

Qy       1000 ProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerPro 1019
                       :::      |||   :::   |||         :::      :::||||||
Db       3163 ---GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAATGTAGAAGTTCTACAAGTCCA 3219

Qy       1020 ArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAla 1039
                 |||   |||:::||||||            ::::::|||
Db       3220 TGTGGCACATCCAAGAGTCCA------------AACAGAGAG------------------ 3249

Qy       1040 GlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPro 1059
                                         |||
Db       3250 ---------------------------TGG------------------------------ 3252
```

```
Qy      1060 ArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPro 1079
                ||||||::::::   :::||||||                     :::      |||
Db      3253 ---GAAGTCCTTCAGCCTGCTCCACAT------------CAATTGATAACTAATCTCCCT 3297

Qy      1080 ProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPro 1099
                |||      ||||||               |||   ||||||      |||   |||
Db      3298 GAAGGCGTTCGGCTTCCG---------------ACAACTCGACCAACCAGGCCACCGCCC 3342

Qy      1100 ProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGlu 1119
             |||   |||   :::            ::::::|||   |||   ||||||   :::
Db      3343 CCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAAAAA---CCATCTTTTATA--- 3396

Qy      1120 ArgGlnIleGlyAlaIleSerGlnGly---------------MetSerValGlnLeuHis 1134
                      |||:::||||||||||||               ::::::
Db      3397 ---ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACTTATTTGACTTCTCATAATCAG 3453

Qy      1135 ValProTyrSerGluHisAlaLysAlaPro---ValGlyProValThrMetGlyLeuPro 1153
                   |||::::::            |||   ||||||   :::::::::|||||||||
Db      3454 GCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTAGGATCTATCTCTCTTGGACTGCCA 3513

Qy      1154 LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSer 1173
                      :::      |||   |||            :::|||||||||:::   |||
Db      3514 CGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTACATCAAGCAGGAAGAATTTTCT 3573

Qy      1174 ProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSer 1193
             ||||||   |||         ||||||   |||   |||      ||||||
Db      3574 CCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC---AGGGCCCAACATGAAGGT 3630

Qy      1194 ValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIlePro 1213
             |||:::||||||||||||   |||::::::      ||||||||||||:::|||   |||
Db      3631 GTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGAAGTATAACTCGGGGAACTCCA 3687

Qy      1214 SerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThr 1233
             ::::::::::::      :::|||      :::   |||||||||||||||   ||||||
Db      3688 ACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGGGGCTCTATCACTCAGGGCACC 3747

Qy      1234 ProAla------------------------AspValLeuTyrLysGlyThrIleThrArg 1245
             ||||||                        :::   |||   ||||||:::|||:::|||
Db      3748 CCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTGGTGAAGGGGTCCATTTCGAGA 3807

Qy      1246 IleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLys 1265
             :::      |||||||||            ::::::|||||||||::::::      |||
Db      3808 ATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGCAGAGAGGAAGCTGCATCCAAA 3861

Qy      1266 GlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSer 1285
             ||||||||||||||||||||||||   ||||||:::|||||||||:::         :::
Db      3862 GGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTGTCATATGATAATATTAAGAAT 3921

Qy      1286 ValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAla 1305
                                  :::|||      |||            ||||||   :::
Db      3922 GCC---------------CGAGAAGGGACTAGGAGTCCAAGAACAGCTCATGAAATCAGT 3966

Qy      1306 AlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer------ 1323
                   ||||||:::|||:::   :::||||||   :::   :::   :::|||
Db      3967 TTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATAAAGCAAGGGATGTCAATGAGG 4023

Qy      1324 ------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArg 1339
                         ||||||   :::|||||||||:::   ||||||:::|||      |||
Db      4024 GAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGCCGAGCATTACCC------AGG 4077

Qy      1340 HisSerProHis---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358
                |||||||||      |||||||||:::      :::   |||||||||   ||||||
Db      4078 GGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTGTCTGGCTCCATAATGCAGGGG 4137

Qy      1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378
                ||||||:::      |||:::   ||||||   |||   :::      |||   :::
Db      4138 ACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT---AAATATCCCAAACAAATT 4194

Qy      1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
             |||||||||                  ||||||   |||            |||
Db      4195 AAAAGGGAAAGT---------------CCTCCCATACGAGCATTTGAAGGTGCCATT--- 4236

Qy      1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
                                  |||   ||||||   ::::::|||   :::   |||:::
Db      4237 ------------------ACCAAAGGAAAACCA---TATGATGGC---ATCACCACCATC 4272

Qy      1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu--------------- 1433
             ||||||   |||||||||||||||||||||||||||:::::::::
Db      4273 AAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAAGATATTTTAACTCAGGAAAGT 4332

Qy      1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
             |||   |||||||||:::      :::   ||||||:::   ||||||||||||:::|||
Db      4333 CGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATAATTGAGGGTTCCATTTCCCAG 4392

Qy      1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
             |||||||||:::|||:::|||         |||         :::   ||||||:::|||
Db      4393 GGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAATCTGCCATCAAACACAATGTC 4449

Qy      1474 ArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
             :::|||||||||      |||   :::            :::   ||||||:::::::::
Db      4450 AAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGAATGCCTCCGCTGGAAATTGTG 4509

Qy      1494 AlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu---------------GluSer 1507
                :::      :::   :::||||||      ||||||               |||:::
Db      4510 CCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAGACC 4569

Qy      1508 LeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaPro 1527
             ::::::||||||      :::      |||||||||   |||:::   |||
Db      4570 GTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCCTCCGTTCTTAGGTCCACA--- 4626

Qy      1528 ValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGly 1547
                   :::   |||      |||   :::   ||||||      |||:::|||
Db      4627 ------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCTGGGATTTATGATGACACCAGT 4680

Qy      1548 Ala---------------ProPheAlaGlyHisLeuProArgGlySerProValThrMet 1562
             |||                  :::         :::   ||||||||||||:::
Db      4681 GCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCCAGAGGCTCACCCATGATGAAC 4740

Qy      1563 ArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAsp 1582
             |||               :::                  :::         :::   :::
Db      4741 AGAACTTCTGATGTTACAATTCCT------------CCTAACAAGTCTACCAATCATGAA 4788

Qy      1583 ArgLysLeuThrSerThrProArgGlu------------IleAlaLysSerProHisSer 1598
             ||||||   |||   ||||||   :::               ||||||||||||
Db      4789 AGGAAATCGACACTGACCCCTACCCAGAGGGAAAGTATCCCAGCGAAGTCTCCAGTGCCT 4848

Qy      1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
                |||               |||      ||||||::::::         ||||||
Db      4849 GGGGTGGACCCTGTCGTGAGCCAC------AGTCCGTTTGATCCCCATCACAGAGGCAGC 4902

Qy      1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
             :::      ::::::|||   ||||||:::|||         ||||||   ::::::|||
Db      4903 ACTGCAGGCGAGGTTTATTGGAGCCACCTGCCCACGCAATTGGATCCA---GCCATGCCT 4959

Qy      1639 ArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaPro 1657
                         ||||||   |||||||||||||||         |||   |||:::|||
Db      4960 TTTCACAGGGCTTTGGATCCTGCAGCGGCTGCTTACCTGTTTCAGAGACAGCTTTCACCA 5019

Qy      1658 AsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAla 1677
                |||   ||||||      |||      |||
Db      5020 ACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC--------------------------- 5052

Qy      1678 AlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHis 1696
             |||:::||||||   ||||||||||||:::|||||||||||||||||||||||||||
Db      5053 GCAATGGAGAACACAAGACAGACAATCTTAAATGATTACATTACCTCACAACAGATGCAA 5112

Qy      1697 HisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArg 1716
                |||                     |||   |||:::   ||||||||||||||||||
Db      5113 GTGAACTTG------------------CGTCCAGATGTGGCCAGAGGACTCTCCCCAAGA 5154

Qy      1717 GluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGln 1736
             |||      |||   |||   |||   |||      ||||||||||||||||||:::
Db      5155 GAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGAGGAATCATTGACCTGACCAAT 5211

Qy      1737 ValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAsp 1756
             :::|||         :::|||||||||         ||||||            ||||||
Db        5212 ATGCCTCCA---ACAATTTAGTGCCTCATCCAGGGGGAACAAGCACTCCTCCCATGGAC 5268

Qy        1757 ArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg---HisSerSerSer 1775
               |||:::   |||:::|||               |||      |||   ::::::|||:::
Db        5269 AGAATCACTTATATTCCTGGTACACAGATTACTTTCCCTCCCAGGCCGTACAACTCTGCT 5328

Qy        1776 ProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGlu 1795
                  :::|||||||||   ||||||||||||                  :::|||:::|||
Db        5329 TCCATGTCTCCAGGACACCCAACACACCTT---------GCAGCTGCTGCAAGTGCTGAG 5379

Qy        1796 ArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeu 1815
               |||||||||:::|||:::||||||::::::|||:::|||||||||         |||
Db        5380 AGGGAACGGAACGGAGCGGAGAAGGAGCGGAGCGGAACGG---------ATTGCT 5430

Qy        1816 ThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSer 1835
                  :::::::::   :::                  |||||||||:::||||||
Db        5431 GCAGCTTCCTCCGACCTC------------TACCTGCGGCCAGGCTCAGAACAG------ 5472

Qy        1836 GlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSer 1855
                                                         ||||||   ||||||
Db        5473 ------------------------------------CCTGGCCGACCTGGCAGTCATGGA 5496

Qy        1856 HisAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGlnGlnArg 1874
               :::            ||||||   |||   |||||||||:::      ||||||||||||
Db        5497 TATGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAGGAGACCATGTTGCAACAGAGA 5553

Qy        1875 ProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLys 1894
               |||||||||         |||            :::||||||   ::::::|||:::
Db        5554 CCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATCACACCTTTGGATCCAACTGCT 5613

Qy        1895 ProThrValLeuArgSerThrSerThrSerSerProVal------------ArgProAla 1910
                        :::                     |||                  ||||||
Db        5614 CAGCTACGAATCATGCCACTGCCTGCTGGGGGCCCTTCAATAAGCCAAGGCCTGCCAGCC 5673

Qy        1911 AlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPro 1930
               :::   :::      |||            |||         :::|||
Db        5674 TCCCGTTACAACACTGCTGCGGATGCC---CTGGCTGCTCTTGTGGATGCTGCAGCTTCT 5730

Qy        1931 ThrLeuMetGluProValLeuLeuProLys------------GluAlaProArgValAla 1946
                              |||         |||            ||||||   |||:::
Db        5731 GCACCCCAGATGGATGTGTCCAAAACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTAGAA 5790

Qy        1947 ArgProGluArgProArgAlaAsp------------------------------------ 1954
                        |||   |||:::
Db        5791 GAAAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAACC 5850

Qy        1955 ------------------------------------ThrGlyHisAlaPheLeuAlaLys 1962
                                                   |||      ||||||   :::
Db        5851 CTGGAGGTGGAGAAGAGATCTGTTCAGTGTTTATACACTTCTTCAGCCTTTCCAAGTGGC 5910

Qy        1963 ProProAlaArgSerGlyLeuGluProAlaSerSer---------ProSerLysGlySer 1979
                  |||               :::|||   ||||||                  |||
Db        5911 AAGCCC---------------CAGCCTCATTCTTCAGTAGTTTATTCTGAGGCTGGGAAA 5955

Qy        1980 GluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAla 1999
               :::      |||      ||||||   |||   :::            ||||||
Db        5956 GATAAAGGGCCT------CCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCAGAGGG 6009

Qy        2000 LysAsn------LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSer 2017
               |||            |||   :::                              ||||||
Db        6010 AAGACTACCATTACTGCAGCTAACTTCATAGACGTGATCATCACCCGGCAAATTGCCTCG 6069

Qy        2018 AlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGlu 2037
                     |||               :::||||||         |||
Db        6070 GACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT--------------- 6114

Qy        2038 LeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSer 2057
                     ||||||      |||                  :::   :::|||   :::|||
Db        6115 ---AGTAGCTTATCTTCTCACAGGTATGAAACACCTAGCGATGCTATTGAGGTGATAAGT 6171

Qy        2058 ProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeu 2077
               |||   |||||||||:::         :::|||                  |||   :::
Db        6172 CCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAGACCTATCAGCCAGAGGTTGTT 6231

Qy        2078 AspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGly 2097
                     :::      |||   :::         :::      ||||||
Db        6232 AAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATATGAAGGACCA------------ 6279

Qy        2098 GlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSer 2117
                              |||   |||   ||||||            |||   |||   |||
Db        6280 ---------------TTACATCACTATCGACCA---------CAGCAGGAATCACCATCT 6315

Qy        2118 ProLeuLeuGlnThrAlaProGly---------------------------ValLysGlyHis 2129
               |||      |||      |||                           |||
Db        6316 CCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGAATGGGGCAAGTGCCCAGGACC 6375

Qy        2130 GlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArg 2149
                  |||::::::|||||||||   ||||||   ::::::||||||||||||:::   |||
Db        6376 CATCGGCTGATCACACTTGCTGATCACATCTGTCAAATTATCACACAAGATTTTGCTAGA 6435

Qy        2150 HisHis---------------ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyr 2164
               :::                  |||||||||         |||      :::
Db        6436 AATCAAGTTTCCTCGCAGACTCCCCAGCAG---------CCTCCTACTTCTACATTCCAG 6486

Qy        2165 SerPheProGlyAla------SerCysProValLeuAspLeuArgArgProProSerAsp 2182
               :::   |||   |||      |||   ||||||         |||         |||:::
Db        6487 AACTCACCTTCTGCTTTGGTATCTACACCTGTG---------AGGACTAAAACATCAAAC 6537

Qy        2183 LeuTyrLeuProProProAsp---------------HisGlyAlaPro---AlaArgGly 2196
                  |||   |||                        |||      |||   :::|||
Db        6538 CGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCATCAAAGACCAGGTTCAAGGGTC 6597

Qy        2197 SerProHis---------------SerGluGlyGlyLysArgSerProGluProAsnLys 2211
               ||||||                  |||   |||   :::            |||   :::
Db        6598 TCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGGCCTGGAAAATCCCCAGAGAGG 6657

Qy        2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
               :::   |||               :::      ||||||:::|||||||||:::
Db        6658 AGTCACGTC---------TCTTCCGAGCCCTACGAGCCCATCTCCCCACCCCAG------ 6702

Qy        2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
                           |||   :::               ||||||   :::   |||   :::
Db        6703 GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTCTTGTCTCAGAGGGGCGCAGAG 6762

Qy        2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
                           |||      ::::::|||||||||:::   |||      |||:::|||
Db        6763 CCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGTATAAGCTACTTGCCTTCATTC 6822

Qy        2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
               |||:::||||||   |||:::   |||   |||||||||||||||||||||||||||
Db        6823 TTCACCAAGCTT---GAAAATACATCACCCATGGTTAAATCAAAGAAGCAGGAGATTTTT 6879

Qy        2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
               :::|||||||||:::            :::   :::         :::||||||||||||
Db        6880 CGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGAACT 6939

Qy        2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
               ||||||||||||:::||||||:::|||   :::|||   ::::::   |||      :::
Db        6940 GAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCAGTTAGCTCTAGAGGCCATTCT 6999

Qy        2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
                     :::   ||||||   |||:::|||||||||   |||||||||||||||||||||
Db        7000 TTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGACATTATCAGGAAGGCTCTCATG 7056

Qy        2352 GlyLysTyrAspGlnTrpGluGlu------------------SerProProLeuSerAla 2365
               |||   :::|||         |||                  |||   |||:::
Db        7057 GGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTCATGTCCCAGCCTATGGGAGTA 7116

Qy        2366 AsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAla 2385
                           |||      |||:::   |||:::               :::|||:::
Db        7117 ---------GTGCCTGGTACTGCCAACACCTCAGTT---------------GTGACCAGT 7152

Qy        2386 AlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGly------LysAla 2403
                  :::   |||   :::            ||||||      ||||||      |||
Db        7153 GGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCATTCAGGAGGAGTTTGCAAACCA 7212
```

```
Qy       2404 LysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro------GlyLeu 2421
              |||:::      :::   :::|||||||||:::|||||||||   |||      |||
Db       7213 AAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCTATACCTGGGCAAGGCTAC 7272

Qy       2422 AlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArg 2441
                       :::||||||   ||||||||||||||||||||||||||||||   :::|||
Db       7273 TTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCAGAAGGGGATTACCATAGG 7332

Qy       2442 ArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrPro 2461
              :::||||||               |||||||||||||||||||||   |||||||||
Db       7333 CAGACGCCA------GGGTGGGCCTGGGAAGACAGGCCCTCTTCAACAGGCTCAACTCAG 7386

Qy       2462 PheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProPro 2481
              ||||||||||||||||||   ||||||:::         :::   ::::::||||||
Db       7387 TTTCCTTATAACCCTCTGACTATGCGGATG---------CTCAGCAGTACTCCACCAACA 7437

Qy       2482 ProGlyLeuProAlaGlySerGly---ProLeuAlaGlyProHis---------HisAla 2497
              |||         |||   |||            |||   ||||||
Db       7438 CCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCTCACCAACAGAACAGGATC 7497

Qy       2498 TrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||:::   ||||||   |||||||||   :::||||||||||||||||||||||||:::
Db       7498 TGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAGACCCTGTCGGATAGTGAT 7557

RESULT 14
AF044209
LOCUS       AF044209                7940 bp    mRNA    linear   PRI 01-SEP-1998
DEFINITION  Homo sapiens nuclear receptor co-repressor N-CoR mRNA, complete
            cds.
ACCESSION   AF044209
VERSION     AF044209.1  GI:3510602
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7940)
  AUTHORS   Wang,J., Hoshino,T., Redner,R.L., Kajigaya,S. and Liu,J.M.
  TITLE     ETO, fusion partner in t(8;21) acute myeloid leukemia, represses
            transcription by interaction with the human N-CoR/mSin3/HDAC1
            complex
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10860-10865 (1998)
  MEDLINE   98393736
   PUBMED   9724795
REFERENCE   2  (bases 1 to 7940)
  AUTHORS   Wang,J. and Liu,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-1998) Hematology Branch, National Heart, Lung and
            Blood Institute, Bldg. 10, Rm. 7C103, 9000 Rockville Pike,
            Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1..7940
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             241..7563
                     /codon_start=1
                     /product="nuclear receptor co-repressor N-CoR"
                     /protein_id="AAC33550.1"
                     /db_xref="GI:3510603"
                     /translation="MSSSGYPPNQGAFSTEQSRYPPHSVQYTFPNTRHQQEFAVPDYR
                     SSHLEVSQASQLLQQQQQQLRRRPSLLSEFHPGSDRPQERRTSYEPFHPGPSPVDHD
                     SLESKRPRLEQVSDSHFQRVSAAVLPLVHPLPEGLRASADAKKDPAFGGKHEAPSSPI
                     SGQPCGDDQNASPSKLSKEELIQSMDRVDREIAKVEQQILKLKKKQQQLEEEAAKPPE
                     PEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKIFEGLGPKVELPLYNQPSDTKVYH
                     ENIKTNQVMRKKLILFFKRRNHARKQREQKICQRYDQLMEAWEKKVDRIENNPRRKAK
                     ESKTREYYEKQFPEIRKQREQQERFQRVGQRGAGLSATIARSEHEISEIIDGLSEQEN
                     NEKQMRQLSVIPPMMFDAEQRRVKFINMNGLMEDPMKVYKDRQFMNVWTDHEKEIFKD
                     KFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKALVRRNYGKRRGRNQQIARP
                     SQEEKVEEKEEDKAEKTEKKEEEKKDEEEKDEKEDSKENTKEKDKIDGTAEETEEREQ
                     ATPRGRKTANSQGRRKGRITRSMTNEAAAASAAAAAATEEPPPPLPPPPEPISTEPVE
                     TSRWTEEEMEVAKKGLVEHGRNWAAIAKMVGTKSEAQCKNFYFNYKRRHNLDNLLQQH
                     KQKTSRKPREERDVSQCESVASTVSAQEDEDIEASNEEENPEDSEVEAVKPSEDSPEN
                     ATSRGNTEPAVELEPTTETAPSTSPSLAVPSTKPAEDESVETQVNDSISAETAEQMDV
                     DQQEHSAEEGSVCDPPPATKADSVDVEVRVPENHASKVEGDNTKERDLDRASEKVEPR
                     DEDLVVAQQINAQRPEPQSDNDSSATCSADEDVDGEPERQRMFPMDSKPSLLNPTGSI
                     LVSSPLKPNPLDLPQLQHRAAVIPPMVSCTPCNIPIGTPVSGYALYQRHIKAMHESAL
                     LEEQRQRQEQIDLECRSSTSPCGTSKSPNREWEVLQPAPHQLITNLPEGVRLPTTRPT
                     RPPPPLIPSSKTTVASEKPSFIMGGSISQGTPGTYLTSHNQASYTQETPKPSVGSISL
                     GLPRQQESAKSATLPYIKQEEFSPRSQNSQPEGLLVRAQHEGVVRGTAGAIQEGSITR
                     GTPTSKISVESIPSLRGSITQGTPALPQTGIPTEALVKGSISRMPIEDSSPEKGREEA
                     ASKGHVIYEGKSGHILSYDNIKNAREGTRSPRTAHEISLKRSYESVEGNIKQGMSMRE
                     SPVSAPLEGLICRALPRGSPHSDLKERTVLSGSIMQGTPRATTESFEDGLKYPKQIKR
                     ESPPIRAFEGAITKGKPYDGITTIKEMGRSIHEIPRQDILTQESRKTPEVVQSTRPII
                     EGSISQGTPIKFDNNSGQSAIKHNVKSLITGPSKLSRGMPPLEIVPENIKVVERGKYE
                     DVKAGETVRSRHTSVVSSGPSVLRSTLHEAPKAQLSPGIYDDTSARRTPVSYQNTMSR
                     GSPMMNRTSDVTIPPNKSTNHERKSTLTPTQRESIPAKSPVPGVDPVVSHSPFDPHHR
                     GSTAGEVYWSHLPTQLDPAMPFHRALDPAAAAYLFQRQLSPTPGYPSQYQLYAMENTR
                     QTILNDYITSQQMQVNLRPDVARGLSPREQPLGLPYPATRGIIDLTNMPPTILVPHPG
                     GTSTPPMDRITYIPGTQITFPPRPYNSASMSPGHPTHLAAAASAERERERERERKERER
                     ERIAAASSDLYLRPGSEQPGRPGSHGYVRSPSPSVRTQETMLQQRPSVFQGTNGTSVI
                     TPLDPTAQLRIMPLPAGGPSISQGLPASRYNTAADALAALVDAAASAPQMDVSKTKES
                     KHEAARLEENLRSRSAAVSEQQQLEQKTLEVEKRSVQCLYTSSAFPSGKPQPHSSVVY
                     SEAGKDKGPPPKSRYEEELRTRGKTTITAANFIDVIITRQIASDKDARERGSQSSDSS
                     SSLSSHRYETPSDAIEVISPASSPAPPQEKLQTYQPEVVKANQAENDPTRQYEGPLHH
                     YRPQQESPSPQQQLPPSSQAEGMGQVPRTHRLITLADHICQIITQDFARNQVSSQTPQ
                     QPPTSTFQNSPSALVSTPVRTKTSNRYSPESQAQSVHHQRPGSRVSPENLVDKSRGSR
                     PGKSPERSHVSSEPYEPISPPQVPVVHEKQDSLLLLSQRGAEPAEQRNDARSPGSISY
                     LPSFFTKLENTSPMVKSKKQEIFRKLNSSGGGDSDMAAAQPGTEIFNLPAVTTSGSVS
                     SRGHSFADPASNLGLEDIIRKALMGSFDDKVEDHGVVMSQPMGVVPGTANTSVVTSGE
                     TRREEGDPSPHSGGVCKPKLISKSNSRKSKSPIPGQGYLGTERPSSVSSVHSEGDYHR
                     QTPGWAWEDRPSSTGSTQFPYNPLTMRMLSSTPPTPIACAPSAVNQAAPHQQNRIWER
                     EPAPLLSAQYETLSDSDD"
ORIGIN

Alignment Scores:
Pred. No.:              1.07e-75        Length:         7940
Score:                  4147.50         Matches:        1099
Percent Similarity:     53.33%          Conservative:   341
Best Local Similarity:  40.70%          Mismatches:     785
Query Match:            31.38%          Indels:         475
DB:                     9               Gaps:           108

US-09-522-753-5 (1-2517) x AF044209 (1-7940)

Qy         16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
              :::   ||||||||||||||||||:::   |||               |||
Db        289 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 348

Qy         36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
              :::      :::   :::|||:::      |||               ||||||   |||
Db        349 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 408

Qy         54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                |||      ||||||||||||||||||||||||||||||
Db        409 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 468

Qy         74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
              ||||||::::::|||   ||||||         |||               :::   |||
Db        469 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 522

Qy         94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                             :::   :::   :::||||||||||||||||||||||||   :::
Db        523 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 582

Qy        114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                 |||                                 ||||||:::   |||   |||
Db        583 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 642

Qy        123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                          |||      |||   |||   |||      ||||||   :::
Db        643 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 690

Qy        143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
              ||||||   |||   |||   ||||||   |||                  :::
Db        691 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 750
```

```
Qy         162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                     |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db         751 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 810

Qy         182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
               ||||||      |||||||||||||||   ||||||||||||||||||||||||||||||
Db         811 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 870

Qy         202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db         871 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 930

Qy         222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
               ||||||||||||:::|||||||||||||||||||||||||||||||||||||||   |||
Db         931 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAAAAAGCAGAAGAAGCT 990

Qy         242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
               |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db         991 CATAAAATTTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 1050

Qy         262 AspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIle 281
               ||||||:::   ||||||||||||||||||   ||||||   ||||||||||||||||||
Db        1051 GATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTGATGAGGAAAAAACTCATT 1110

Qy         282 LeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArg 301
               |||:::||||||||||||||||||||||||||||||   :::||||||   |||||||||
Db        1111 TTATTTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAACAAAAAATCTGCCAGCGT 1170

Qy         302 TyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArg 321
               |||||||||||||||||||||   ||||||||||||:::|||||||||||||||||||||
Db        1171 TATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGAATAGAAAATAATCCTCGG 1230

Qy         322 ArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArg 341
               |||:::|||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        1231 AGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAGCAGTTTCCAGAAATTCGA 1290

Qy         342 LysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeu 361
               ||||||||||||   |||||||||   |||   ||||||||||||||||||:::||||||
Db        1291 AAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGGCAGAGGGGAGCTGGTCTT 1347

Qy         362 SerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGlu 381
               |||   :::   ||||||||||||||||||:::|||||||||||||||||||||||||||
Db        1348 TCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATTATTGATGGGCTCTCTGAG 1407

Qy         382 GlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAsp 401
               |||||||||   |||||||||||||||||||||:::|||||||||||||||::::::|||
Db        1408 CAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGATTCCACCTATGATGTTTGAT 1467

Qy         402 AlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLys 421
               |||:::|||:::|||:::|||||||||||||||||||||||||||   ||||||||||||
Db        1468 GCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCTTATGGAGGACCCTATGAAA 1527

Qy         422 ValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArg 441
               ||||||||||||||||||   ||||||:::|||::::::   |||||||||   |||:::
Db        1528 GTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCATGAAAAGGAGATCTTTAAG 1587

Qy         442 GluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLys 461
               :::||||||:::|||||||||||||||||||||||||||||||||:::||||||||||||
Db        1588 GACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCATCATACTTGGAGAGGAAG 1647

Qy         462 ThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSer 481
               :::|||   :::|||||||||||||||||||||||||||||||||||||||||||||:::
Db        1648 AGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAAAATGAGAATTATAAAGCC 1707

Qy         482 LeuValArgArgSerTyr---ArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
               ||||||||||||:::|||   :::|||||||||::::::||||||      :::
Db        1708 CTCGTCAGAAGGAATTATGGGAAACGCAGAGGCAGAAACCAGCAAATTGCTCGACCCTCG 1767

Qy         501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
               |||:::::::::   :::::::::::::::         ::::::   :::   ||||||
Db        1768 CAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAAGCAGAAAAAACAGAAAAA 1821

Qy         521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                  |||:::||||||:::   |||   ||||||:::|||::::::         ||||||
Db        1822 AAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAAGAAGACTCCAAAGAAAAT 1881

Qy         541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                  ||||||         |||:::|||   |||   |||:::   |||:::   :::|||
Db        1882 ACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA---GAAGAAACTGAGGAA 1929

Qy         561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
               :::|||            :::|||||||||||||||||||||||||||||||||||||||
Db        1930 AGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGTCAGGGCCGCCGTAAGGGC 1989

Qy         581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
               ||||||||||||||||||   |||||||||   :::      |||
Db        1990 CGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGTGCTGCAGCCGCAGCGGCT 2049

Qy         601 Ser------------------------------AlaGluLeuAlaSerMetGluLeuAsn 610
               :::                                 |||      |||   |||
Db        2050 ACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCCATTTCTACAGAGCCTGTG 2109

Qy         611 GluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLeuGluHis 630
               |||:::|||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db        2110 GAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAAAAAGGTCTAGTAGAACAT 2169

Qy         631 GlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerGlnCysLys 650
               ||||||||||||:::|||||||||:::|||||||||:::|||:::   :::|||||||||
Db        2170 GGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAAAGTGAAGCTCAATGTAAA 2229

Qy         651 AsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeuGlnGlnHisLys 670
               |||||||||||||||||||||:::|||   |||||||||   :::|||||||||||||||
Db        2230 AACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAACCTCTTACAGCAGCATAAA 2289

Qy         671 LeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAlaAla 690
                  |||      ::::::            |||         :::      :::   |||
Db        2290 CAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCTCAATGTGAAAGTGTCGCT 2349

Qy         691 SerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGlyVal 710
               |||      :::|||               |||||||||::::::|||||||||
Db        2350 TCCACTGTTTCTGCT------------CAGGAGGATGAAGATATTGAAGCCTCC------ 2391

Qy         711 SerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsnGlu 730
                     ||||||||||||      |||::::::|||
Db        2392 ------AATGAAGAAGAAAATCCAGAAGACAGCGAA------------------------ 2421

Qy         731 ValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSerAspThrGluSer 750

Db        2421 ------------------------------------------------------------ 2421

Qy         751 IleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGlyProLysProPro 770
                                 ||||||   |||
Db        2422 ---------------GTTGAAGCTGTCAAG------------------------------ 2436

Qy         771 AlaThrLeuGlyAlaAspGlyProProProGlyProProThrProProArgArgThrSer 790
                                          |||         :::|||         ||||||
Db        2437 ---------------------------CCCAGCGAGGACAGTCCTGAAAATGCTACTTCT 2469

Qy         791 ArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProPro 810
               |||         ||||||            |||            ||||||
Db        2470 CGAGGAAACACAGAACCTGCGGTTGAGCTTGAG------------CCCACCACGGAAACT 2517

Qy         811 AlaPro---ProSerProSer---AlaProProProValValProLysGluGluLysGlu 828
               ||||||      |||||||||   |||   |||         |||   |||::::::
Db        2518 GCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACAAAACCAGCTGAAGATGAAAGT 2577

Qy         829 GluGluThrAlaAlaAlaProProVal--------------------------------- 837
                  ||||||               :::
Db        2578 GTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACAGCAGAGCAGATGGATGTAGAT 2637

Qy         838 ------------------GluGluGlyGluGluGlnLysProProAlaAlaGluGluLeu 851
                                 |||||||||            ||||||   |||   :::
Db        2638 CAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGATCCCCCACCCGCTACCAAA--- 2694

Qy         852 AlaValAspThrGlyLysAlaGluGluProValLysSerGluCysThrGluGluAlaGlu 871
```

```
                    |||:::         |||      |||                  :::   |||
Db       2695 ---GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAAAACCATGCATCTAAAGTTGAA 2751

Qy        872 GluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeu 891
                          |||   :::|||   :::   |||
Db       2752 GGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC------------------------ 2787

Qy        892 LysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLysSerSerGlyAla 911
                 :::||||||   |||                        |||:::         |||
Db       2788 ---AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTGGTAGCTCAGCAAATAAATGCC 2844

Qy        912 ------------ProGln---AspSerAspSerSerAlaThrCysSerAlaAspGluVal 926
                          ||||||   |||:::||||||||||||||||||||||||||||||
Db       2845 CAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCCACGTGCAGCGCTGATGAG--- 2901

Qy        927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro---------ArgProSer 943
                 :::   :::|||      :::   :::   :::   |||         :::||||||
Db       2902 ---GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTTCCTATGGACTCAAAGCCTTCA 2958

Qy        944 LeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys------ProLeu 961
              ||||||   |||||||||               :::||||||   |||      ||||||
Db       2959 CTGTTAAACCCCACTGGATCTATACTCGTC---TCATCTCCGTTAAAACCAAATCCACTG 3015

Qy        962 AspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIle--------------- 976
              ||||||   ||||||:::   |||||||||   |||||||||:::
Db       3016 GATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCACCAATGGTATCCTGCACCCCA 3075

Qy        977 ---------------------------------------------------GlnValThr 979
                                                                    :::
Db       3076 TGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCTCTCTACCAGCGACACATTAAA 3135

Qy        980 LysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaPro 999
                 :::||||||            :::
Db       3136 GCAATGCATGAGTCAGCACTCCTGGAG--------------------------------- 3162

Qy       1000 ProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerPro 1019
                       :::      |||   :::   |||         :::      :::||||||
Db       3163 ---GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAATGTAGAAGTTCTACAAGTCCA 3219

Qy       1020 ArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAla 1039
                 |||   |||:::||||||            ::::::|||
Db       3220 TGTGGCACATCCAAGAGTCCA------------AACAGAGAG------------------ 3249

Qy       1040 GlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProPro 1059
                                         |||
Db       3250 ---------------------------TGG------------------------------ 3252

Qy       1060 ArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPro 1079
                 ||||||::::::   :::||||||                     :::      |||
Db       3253 ---GAAGTCCTTCAGCCTGCTCCACAT------------CAATTGATAACTAATCTCCCT 3297

Qy       1080 ProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPro 1099
                 |||      ||||||               |||   ||||||      |||   |||
Db       3298 GAAGGCGTTCGGCTTCCG---------------ACAACTCGACCAACCAGGCCACCGCCC 3342

Qy       1100 ProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGlu 1119
              |||   |||   :::            ::::::|||   |||   ||||||   :::
Db       3343 CCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAAAAA---CCATCTTTTATA--- 3396

Qy       1120 ArgGlnIleGlyAlaIleSerGlnGly---------------MetSerValGlnLeuHis 1134
                       |||:::||||||||||||               ::::::
Db       3397 ---ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACTTATTTGACTTCTCATAATCAG 3453

Qy       1135 ValProTyrSerGluHisAlaLysAlaPro---ValGlyProValThrMetGlyLeuPro 1153
                    |||::::::            |||   ||||||   :::::::::|||||||||
Db       3454 GCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTAGGATCTATCTCTCTTGGACTGCCA 3513

Qy       1154 LeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSer 1173
                       :::      |||   |||            :::|||||||||:::   |||
Db       3514 CGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTACATCAAGCAGGAAGAATTTTCT 3573

Qy       1174 ProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSer 1193
              ||||||   |||         ||||||   |||   |||      ||||||

Db       3574 CCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC---AGGGCCCAACATGAAGGT 3630

Qy       1194 ValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIlePro 1213
              |||:::||||||||||||   |||::::::      ||||||||||||:::|||   |||
Db       3631 GTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGAAGTATAACTCGGGGAACTCCA 3687

Qy       1214 SerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThr 1233
              ::::::::::::      :::|||      :::   |||||||||||||||   ||||||
Db       3688 ACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGGGGCTCTATCACTCAGGGCACC 3747

Qy       1234 ProAla------------------------AspValLeuTyrLysGlyThrIleThrArg 1245
              ||||||                        :::   |||   ||||||:::|||:::|||
Db       3748 CCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTGGTGAAGGGGTCCATTTCGAGA 3807

Qy       1246 IleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLys 1265
              :::      |||||||||            ::::::|||||||||::::::      |||
Db       3808 ATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGCAGAGAGGAAGCTGCATCCAAA 3861

Qy       1266 GlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSer 1285
              ||||||||||||||||||||||||   ||||||:::|||||||||:::         :::
Db       3862 GGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTGTCATATGATAATATTAAGAAT 3921

Qy       1286 ValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAla 1305
                                   :::|||      |||            ||||||   :::
Db       3922 GCC---------------CGAGAAGGACTAGGAGTCCAAGAACAGCTCATGAAATCAGT 3966

Qy       1306 AlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSer------ 1323
                    ||||||:::|||:::   :::||||||   :::   :::   :::|||
Db       3967 TTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATAAAGCAAGGGATGTCAATGAGG 4023

Qy       1324 ------------SerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArg 1339
                          ||||||   :::|||||||||:::   ||||||:::|||      |||
Db       4024 GAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGCCGAGCATTACCC------AGG 4077

Qy       1340 HisSerProHis---HisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358
                 |||||||||      |||||||||:::      :::   |||||||||   ||||||
Db       4078 GGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTGTCTGGCTCCATAATGCAGGGG 4137

Qy       1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378
                 ||||||:::      |||:::   ||||||   |||   :::      |||   :::
Db       4138 ACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT---AAATATCCCAAACAAATT 4194

Qy       1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398
              |||||||||                  ||||||   |||            |||
Db       4195 AAAAGGGAAAGT---------------CCTCCCATACGAGCATTTGAAGGTGCCATT--- 4236

Qy       1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418
                                   |||   ||||||   ::::::|||   :::   |||:::
Db       4237 ------------------ACCAAAGGAAAACCA---TATGATGGC---ATCACCACCATC 4272

Qy       1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeu--------------- 1433
              ||||||   |||||||||||||||||||||||||||:::::::::
Db       4273 AAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAAGATATTTAACTCAGGAAAGT 4332

Qy       1434 ArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGln 1453
              |||   |||||||||:::      :::   ||||||:::   ||||||||||||:::|||
Db       4333 CGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATAATTGAGGGTTCCATTTCCCAG 4392

Qy       1454 GlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspVal 1473
              |||||||||:::|||:::|||         |||         :::   ||||||:::|||
Db       4393 GGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAATCTGCCATCAAACACAATGTC 4449

Qy       1474 ArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMet 1493
              :::|||||||||      |||   :::            :::   ||||||:::::::::
Db       4450 AAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGAATGCCTCCGCTGGAAATTGTG 4509

Qy       1494 AlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu---------------GluSer 1507
                 :::      :::   :::||||||      ||||||               |||:::
Db       4510 CCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAGGATGTGAAAGCAGGCGAGACC 4569

Qy       1508 LeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaPro 1527
              ::::::||||||      :::      |||||||||   |||:::   |||
Db       4570 GTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCCTCCGTTCTTAGGTCCACA--- 4626
```

```
Qy      1528 ValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGly 1547
                   :::   |||      |||   :::   ||||||      |||:::|||
Db      4627 ------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCTGGGATTTATGATGACACCAGT 4680

Qy      1548 Ala---------------ProPheAlaGlyHisLeuProArgGlySerProValThrMet 1562
             |||                  :::         :::   ||||||||||||:::
Db      4681 GCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCCAGAGGCTCACCCATGATGAAC 4740

Qy      1563 ArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAsp 1582
             |||               :::                  :::         :::   :::
Db      4741 AGAACTTCTGATGTTACAATTCCT------------CCTAACAAGTCTACCAATCATGAA 4788

Qy      1583 ArgLysLeuThrSerThrProArgGlu------------IleAlaLysSerProHisSer 1598
             ||||||   |||   ||||||   :::               ||||||||||||
Db      4789 AGGAAATCGACACTGACCCCTACCCAGAGGGAAAGTATCCCAGCGAAGTCTCCAGTGCCT 4848

Qy      1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
                |||               |||      ||||||::::::         ||||||
Db      4849 GGGGTGGACCCTGTCGTGAGCCAC------AGTCCGTTTGATCCCCATCACAGAGGCAGC 4902

Qy      1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
             :::      ::::::|||   ||||||:::|||         ||||||   ::::::|||
Db      4903 ACTGCAGGCGAGGTTTATTGGAGCCACCTGCCCACGCAATTGGATCCA---GCCATGCCT 4959

Qy      1639 ArgGlyIleProLeuAsp---AlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaPro 1657
                         ||||||   |||||||||||||||         |||   |||:::|||
Db      4960 TTTCACAGGGCTTTGGATCCTGCAGCGGCTGCTTACCTGTTTCAGAGACAGCTTTCACCA 5019

Qy      1658 AsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAla 1677
                |||   ||||||      |||      |||
Db      5020 ACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC--------------------------- 5052

Qy      1678 AlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHis 1696
             |||:::||||||   ||||||||||||:::|||||||||||||||||||||||||||
Db      5053 GCAATGGAGAACACAAGACAGACAATCTTAAATGATTACATTACCTCACAACAGATGCAA 5112

Qy      1697 HisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArg 1716
                |||                     |||   |||:::   ||||||||||||||||||
Db      5113 GTGAACTTG---------------------CGTCCAGATGTGGCCAGAGGACTCTCCCCAAGA 5154

Qy      1717 GluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGln 1736
             |||      |||   |||   |||   |||      ||||||||||||||||||:::
Db      5155 GAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGAGGAATCATTGACCTGACCAAT 5211

Qy      1737 ValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAsp 1756
             :::|||         :::|||||||||         ||||||            ||||||
Db      5212 ATGCCTCCA---ACAATTTTAGTGCCTCATCCAGGGGGAACAAGCACTCCTCCCATGGAC 5268

Qy      1757 ArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArg---HisSerSerSer 1775
             |||:::   |||:::|||               |||      |||   ::::::|||:::
Db      5269 AGAATCACTTATATTCCTGGTACACAGATTACTTTCCCTCCCAGGCCGTACAACTCTGCT 5328

Qy      1776 ProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGlu 1795
                :::|||||||||   ||||||||||||                  :::|||:::|||
Db      5329 TCCATGTCTCCAGGACACCCAACACACCTT---------GCAGCTGCTGCAAGTGCTGAG 5379

Qy      1796 ArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeu 1815
             |||||||||:::|||:::||||||::::::|||:::|||||||||         |||
Db      5380 AGGGAACGGGAACGGGAGCGGGAGAAGGAGCGGGAGCGGGAACGG---------ATTGCT 5430

Qy      1816 ThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSer 1835
                :::::::::   :::                  |||||||||:::||||||
Db      5431 GCAGCTTCCTCCGACCTC------------TACCTGCGGCCAGGCTCAGAACAG------ 5472

Qy      1836 GlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSer 1855
                                                       ||||||   ||||||
Db      5473 ------------------------------------CCTGGCCGACCTGGCAGTCATGGA 5496

Qy      1856 HisAlaHisGlnHisSerProIleSerProArgThrGlnAspAla---LeuGlnGlnArg 1874
             :::            ||||||   |||   |||||||||:::      ||||||||||||
Db      5497 TATGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAGGAGACCATGTTGCAACAGAGA 5553
```

```
Qy      1875 ProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLys 1894
             |||||||||         |||            :::||||||   ::::::|||:::
Db      5554 CCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATCACACCTTTGGATCCAACTGCT 5613

Qy      1895 ProThrValLeuArgSerThrSerThrSerSerProVal------------ArgProAla 1910
                      :::                     |||                  ||||||
Db      5614 CAGCTACGAATCATGCCACTGCCTGCTGGGGGCCCTTCAATAAGCCAAGGCCTGCCAGCC 5673

Qy      1911 AlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPro 1930
             :::   :::      |||            |||         :::|||
Db      5674 TCCCGTTACAACACTGCTGCGGATGCC---CTGGCTGCTCTTGTGGATGCTGCAGCTTCT 5730

Qy      1931 ThrLeuMetGluProValLeuLeuProLys------------GluAlaProArgValAla 1946
                            |||         |||            ||||||   |||:::
Db      5731 GCACCCCAGATGGATGTGTCCAAAACAAAAGAGAGTAAGCATGAAGCTGCCAGGTTAGAA 5790

Qy      1947 ArgProGluArgProArgAlaAsp------------------------------------ 1954
                      |||   |||:::
Db      5791 GAAAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAGCAGCAGCTAGAGCAGAAAACC 5850

Qy      1955 ------------------------------------ThrGlyHisAlaPheLeuAlaLys 1962
                                                 |||      ||||||   :::
Db      5851 CTGGAGGTGGAGAAGAGATCTGTTCAGTGTTTATACACTTCTTCAGCCTTTCCAAGTGGC 5910

Qy      1963 ProProAlaArgSerGlyLeuGluProAlaSerSer---------ProSerLysGlySer 1979
                |||               :::|||   ||||||                  |||
Db      5911 AAGCCC---------------CAGCCTCATTCTTCAGTAGTTTATTCTGAGGCTGGGAAA 5955

Qy      1980 GluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAla 1999
             :::      |||      ||||||   |||   :::            ||||||
Db      5956 GATAAAGGGCCT------CCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCAGAGGG 6009

Qy      2000 LysAsn------LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSer 2017
             |||            |||   :::                              ||||||
Db      6010 AAGACTACCATTACTGCAGCTAACTTCATAGACGTGATCATCACCCGGCAAATTGCCTCG 6069

Qy      2018 AlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGlu 2037
                   |||               :::||||||         |||
Db      6070 GACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT--------------- 6114

Qy      2038 LeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSer 2057
                   ||||||      |||                  :::   :::|||   :::|||
Db      6115 ---AGTAGCTTATCTTCTCACAGGTATGAAACACCTAGCGATGCTATTGAGGTGATAAGT 6171

Qy      2058 ProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeu 2077
             |||   |||||||||:::         :::|||                  |||   :::
Db      6172 CCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAGACCTATCAGCCAGAGGTTGTT 6231

Qy      2078 AspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGly 2097
                   :::      |||   :::         :::      ||||||
Db      6232 AAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATATGAAGGACCA------------ 6279

Qy      2098 GlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSer 2117
                            |||   |||   ||||||            |||   |||   |||
Db      6280 ---------------TTACATCACTATCGACCA---------CAGCAGGAATCACCATCT 6315

Qy      2118 ProLeuLeuGlnThrAlaProGly------------------------ValLysGlyHis 2129
             |||      |||      |||                           |||
Db      6316 CCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGAATGGGGCAAGTGCCCAGGACC 6375

Qy      2130 GlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArg 2149
                |||::::::|||||||||   ||||||   ::::::||||||||||||:::   |||
Db      6376 CATCGGCTGATCACACTTGCTGATCACATCTGTCAAATTATCACACAAGATTTTGCTAGA 6435

Qy      2150 HisHis---------------ProGlnGlnLeuSerAlaProLeuProAlaProLeuTyr 2164
             :::                  |||||||||         |||      :::
Db      6436 AATCAAGTTTCCTCGCAGACTCCCCAGCAG---------CCTCCTACTTCTACATTCCAG 6486

Qy      2165 SerPheProGlyAla------SerCysProValLeuAspLeuArgArgProProSerAsp 2182
             :::   |||   |||      |||   ||||||         |||         |||:::
Db      6487 AACTCACCTTCTGCTTTGGTATCTACACCTGTG---------AGGACTAAAACATCAAAC 6537

Qy      2183 LeuTyrLeuProProProAsp---------------HisGlyAlaPro---AlaArgGly 2196
```

```
                 |||   |||                        |||      |||   :::|||
Db       6538 CGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCATCAAAGACCAGGTTCAAGGGTC 6597

Qy       2197 SerProHis---------------SerGluGlyGlyLysArgSerProGluProAsnLys 2211
              ||||||                  |||   |||   :::            |||   :::
Db       6598 TCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGGCCTGGAAAATCCCCAGAGAGG 6657

Qy       2212 ThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGlyMet 2231
              :::   |||               :::      ||||||:::|||||||||:::
Db       6658 AGTCACGTC---------TCTTCCGAGCCCTACGAGCCCATCTCCCCACCCCAG------ 6702

Qy       2232 ThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGluGln 2251
                          |||   :::               ||||||   :::   |||   :::
Db       6703 GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTCTTGTCTCAGAGGGGCGCAGAG 6762

Qy       2252 ThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhe 2271
                          |||      ::::::|||||||||:::   |||      |||:::|||
Db       6763 CCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGTATAAGCTACTTGCCTTCATTC 6822

Qy       2272 PheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsn 2291
              |||:::||||||   |||:::   |||   |||||||||||||||||||||||||||
Db       6823 TTCACCAAGCTT---GAAAATACATCACCCATGGTTAAATCAAAGAAGCAGGAGATTTTT 6879

Qy       2292 LysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGlyThr 2311
              :::|||||||||:::            :::   :::         :::||||||||||||
Db       6880 CGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATGGCAGCTGCTCAGCCAGGAACT 6939

Qy       2312 GluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGlnAla 2331
              ||||||||||||:::||||||:::|||   :::|||   ::::::   |||      :::
Db       6940 GAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCAGTTAGCTCTAGAGGCCATTCT 6999

Qy       2332 ValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMet 2351
                    :::   ||||||   |||:::|||||||||   |||||||||||||||||||||
Db       7000 TTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGACATTATCAGGAAGCTCTCATG 7056

Qy       2352 GlyLysTyrAspGlnTrpGluGlu------------------SerProProLeuSerAla 2365
              |||   :::|||         |||                  |||   |||:::
Db       7057 GGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTCATGTCCCAGCCTATGGGAGTA 7116

Qy       2366 AsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAla 2385
                          |||      |||:::   |||:::               :::|||:::
Db       7117 ---------GTGCCTGGTACTGCCAACACCTCAGTT---------------GTGACCAGT 7152

Qy       2386 AlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGly------LysAla 2403
                 :::   |||   :::            ||||||      ||||||      |||
Db       7153 GGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCATTCAGGAGGAGTTTGCAAACCA 7212

Qy       2404 LysValSerGlyArgProSerSerArgLysAlaLysSerProAlaPro------GlyLeu 2421
              |||:::      :::   :::|||||||||:::|||||||||   |||      |||
Db       7213 AAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCTATACCTGGGCAAGGCTAC 7272

Qy       2422 AlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArg 2441
                       :::||||||   ||||||||||||||||||||||||||||||   :::|||
Db       7273 TTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCAGAAGGGGATTACCATAGG 7332

Qy       2442 ArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrPro 2461
              :::||||||               |||||||||||||||||||||   |||||||||
Db       7333 CAGACGCCA------GGGTGGGCCTGGGAAGACAGGCCCTCTTCAACAGGCTCAACTCAG 7386

Qy       2462 PheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyValMetAlaSerProProPro 2481
              ||||||||||||||||||   ||||||:::         :::   ::::::||||||
Db       7387 TTTCCTTATAACCCTCTGACTATGCGGATG---------CTCAGCAGTACTCCACCAACA 7437

Qy       2482 ProGlyLeuProAlaGlySerGly---ProLeuAlaGlyProHis---------HisAla 2497
              |||         |||   |||            |||   ||||||
Db       7438 CCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCTCACCAACAGAACAGGATC 7497

Qy       2498 TrpAspGluGluProLysProLeuLeuCysSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||:::   ||||||   |||||||||   :::||||||||||||||||||||||||:::
Db       7498 TGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAGACCCTGTCGGATAGTGAT 7557

RESULT 15
AF087856
```

```
LOCUS       AF087856                8018 bp    mRNA    linear   PRI 23-JUL-2003
DEFINITION  Homo sapiens nuclear receptor co-repressor mRNA, complete cds.
ACCESSION   AF087856
VERSION     AF087856.1  GI:33150575
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 8018)
  AUTHORS   Tu,Q., Yu,L., Dai,F., Yue,P., Huang,H. and Zhao,S.Y.
  TITLE     Cloning and charactetization of a new human cDNA homologous to
            murine nuclear receptor co-repressor mRNA
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 8018)
  AUTHORS   Zhang,M., Yu,L., Zhou,Y., Hu,P.R., Xin,Y.R. and Zhao,S.Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-AUG-1998) Lab of Human Gene Research, Institute of
            Genetics, Fudan University, No. 220 Handan Rd., Shanghai 200433,
            P.R. China
FEATURES             Location/Qualifiers
     source          1. .8018
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS             85. .7116
                     /note="nrcr"
                     /codon_start=1
                     /product="nuclear receptor co-repressor"
                     /protein_id="AAP97166.1"
                     /db_xref="GI:33150576"
                     /translation="MSSSVYPPNQGAFSTEQSRSPPHSVKYTFPSTHHQQDPAFGGKH
                     EAPSSPISGQPCGDDQNASPSKLSKEELIQSMDRVDREIAKVEQQILKLKKKQQQLEE
                     EAAKPPEPEKPVSPPPVEQKHRSIVQIIYDENRKKAEEAHKIFEGLGPKVELPLYNQP
                     SDTKVYHENIKTNQVMRKKLILFFKRRNHARKQREQKICQRYDQLMEAWEKKVDRIEN
                     NPRRKAKESKTREYYEKQFPEIRKQREQQERFQRVGQRGAGLSATIARSEHEISEIID
                     GLSEQESNEKQMRQLSVIPPMMFDAEQRRVKFINMNGLMEDPMKVYKDRQFMNVWTDH
                     EKEIFKDKFIQHPKNFGLIASYLERKSVPDCVLYYYLTKKNENYKALVRRNYGKRRGR
                     NQQIARPSQEEKVEEKEEDKAEKTEKKEEEKKDEEEKDEKEDSKENTKEKDKIDGTAE
                     ETEEREQATPRGRKTANSQGRRKGRITRSMTNEAAAASAAAAAATEEPPPPLPPPPEP
                     ISTEPVETSRWTEEEMEVAKKGLVEHGRNWAAIAKMVGTKSEAQCKNFYFNYKRRHNL
                     DNLLQQHKQKTSRKPREERDVSQCESVASTVSAQEDEDIEASNEEENPEDSEGAENSS
                     DTESAPSPSPVEAVKPSEDSPENATSRGNTEPAVELEPTTETAPSTSPSLAVPSTKPA
                     EDESVETQVNDSISAETAEQMDVDQQEHSAEEGSVCDPPPATKADSVDVEVRVPENHA
                     SKVEGDNTKERDLDRASEKVEPRDEDLVVAQQINAQRPEPQSDNDSSATCSADEDVDG
                     EPERQRMFPMDSKPSLLNPTGSILVSSPLKPNPLDLPQLQHRAAVIPPMVSCTPCNIP
                     IGTPVSGYALYQRHIKAMHESALLEEQRQRQEQIDLECRSSTSPCGTSKSPNREWEVL
                     QPAPHQVITNLPEGVRLPTTRPTRPPPPLIPSSKTTVASEKPSFIMGGSISQGTPGTY
                     LTSHNQASYTQETPKPSVGSISLGLPRQQESAKSATLPYIKQEEFSPRSQNSQPEGLL
                     VRAQHEGVVRGTAGAIQEGSITRGTPTSKISVESIPSLRGSITQGTPALPQTGIPTEA
                     LVKGSISRMPIEDSSPEKGREEAASKGHVIYEGKSGHILSYDNIKNAREGTRSPRTAH
                     EISLKRSYESVEGNIKQGMSMRESPVSAPLEGLICRALPRGSPHSDLKERTVLSGSIM
                     QGTPRATTESFEDGLKYPKQIKRESPPIRAFEGAITKGKPYDGITTIKEMGRSIHEIP
                     RQDILTQESRKTPEVVQSTRPIIEGSISQGTPIKFDNNSGQSAIKHNVKSLITGPSKL
                     SRGMPPLEIVPENIKVVERGKYEDVKAGETVRSRHTSVVSSGPSVLRSTLHEAPKAQL
                     SPGIYDDTSARRTPVSYQNTMSRGSPMMNRTSDVSSNKSTNHERKSTLTPTQRESIPA
                     KSPVPGVDPVVSHSPFDPHHRGSTAGEVYRSHLPTHLDPAMPFHRALDPAAAYLFQRQ
                     LSPTPGYPSQYQLYAMENTRQTILNDYITSQQMQVNLRPDVARGLSPREQPLGLPYPA
                     TRGIIDLTNMPPTILVPHPGGTSTPPMDRITYIPGTQITFPPRPYNSASMSPGHPTHL
                     AAAASAERERERERERKERERERIAAASSDLYLRPGSEQPGRPGSHGYVRSPSPSVRTQ
                     ETMLQQRPSVFQGTNGTSVITPLDPTAQLRIMPLPAGGPSISQGLPASRYNTAADALA
                     ALVDAAASAPQMDVSKTKESKHEAARLEENLRSRSAAVSEQQQLEQKTLEVEKRSVQC
                     LYTSSAFPSGKPQPHSSVVYSEAGKDKGPPPKSRYEEELRTRGKTTITAANFIDVIIT
                     RQIASDKDARERGSQSSDSSSSLSSHRYETPSDAIEVISPASSPAPPQEKLQTYQPEV
                     VKANQAENDPTRQYEGPLHHYRPQQESPSPQQQLPPSSQAEGMGQVPRTHRLITLADH
                     ICQIITQDFARNQVSSQTPQQPPTSTFQNSPSALVSTPVRTKTSNRYSPESQAQSVHH
                     QRPGSRVSPENLVDKSRGSRPGKSPERSHVSSEPYEPISPPQVPVVHEKQDSLLLLSQ
                     RGAEPAEQRNDARSPGSISYLPSFFTKLENTSPMVKSKKQEIFRKLNSSGGGDSDMAA
                     AQPGTEIFNLPAVTTSGSVSSRGHSFADPASNLGLEDIIRKALMGSFDDKVEDHGVVM
                     SQPMGVVPGTANTSVVTSGETRREEGDPSPHSGVCKPKLISKSNSRKSKSPIPGQGYL
                     GTERPSSVSSVHSEGDYHRQTPGWAWEDRPSSTGSTQFPYNPLTMRMLSSTPPTPIAC
                     APSAVNQAAPHQQNRIWEREPAPLLSAQYETLSDSDD"
     polyA_signal    7984. .7989
```

```
ORIGIN

Alignment Scores:
Pred. No.:              1.39e-73        Length:         8018
Score:                  4043.00         Matches:        1061
Percent Similarity:     54.21%          Conservative:   330
Best Local Similarity:  41.35%          Mismatches:     731
Query Match:            30.59%          Indels:         444
DB:                     9               Gaps:           104

US-09-522-753-5 (1-2517) x AF087856 (1-8018)

Qy         137 LysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProProHis 155
               :::|||   :::      ||||||   |||   |||   ||||||   |||
Db         190 CAGGATCCAGCATTCGGAGGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCA 249

Qy         156 ThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsn 175
                        :::            |||   :::||||||||||||||||||||||||:::
Db         250 TGTGGAGATGATCAAAATGCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGT 309

Qy         176 MetAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLys 195
               ||||||||||||||||||||||||      |||||||||||||||   ||||||||||||
Db         310 ATGGATCGTGTAGATCGAGAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAG 369

Qy         196 LysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSer 215
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         370 AAACAACAACAGCTTGAAGAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCC 429

Qy         216 ProProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArg 235
               |||||||||:::|||   ||||||||||||:::|||||||||||||||||||||||||||
Db         430 CCTCCTCCTGTGGAGCAGAAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGG 489

Qy         236 LysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuPro 255
               ||||||||||||   ||||||:::|||   |||||||||||||||:::||||||||||||
Db         490 AAAAAAGCAGAAGAAGCTCATAAAATTTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCA 549

Qy         256 LeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAla 275
               ||||||||||||||||||||||||:::   ||||||||||||||||||   ||||||
Db         550 CTGTATAACCAGCCATCAGATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGTG 609

Qy         276 MetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLys 295
               |||||||||||||||||||||:::||||||||||||||||||||||||||||||   :::
Db         610 ATGAGGAAAAAACTCATTTTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGAA 669

Qy         296 GlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArg 315
               ||||||   ||||||||||||||||||||||||||||||   ||||||||||||:::|||
Db         670 CAAAAAATCTGCCAGCGTTATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAGA 729

Qy         316 IleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLys 335
               |||||||||||||||||||||:::|||||||||||||||   ||||||||||||||||||
Db         730 ATAGAAAATAATCCTCGGAGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAAG 789

Qy         336 GlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGly 355
               ||||||||||||||||||||||||||||||   |||||||||   |||   |||||||||
Db         790 CAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGGG 846

Qy         356 GlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIle 375
               |||||||||:::|||||||||   :::   ||||||||||||||||||:::|||||||||
Db         847 CAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAATT 906

Qy         376 IleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIle 395
               ||||||||||||||||||||||||:::   |||||||||||||||||||||:::||||||
Db         907 ATTGATGGGCTCTCTGAGCAGGAAAGTAATGAGAAACAAATGCGGCAGCTCTCTGTGATT 966

Qy         396 ProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeu 415
               |||||||||::::::||||||:::|||:::|||:::||||||||||||||||||||||||
Db         967 CCACCTATGATGTTTGATGCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCTT 1026

Qy         416 MetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGln 435
               |||   ||||||||||||||||||||||||||||||   ||||||:::|||::::::
Db        1027 ATGGAGGACCCTATGAAAGTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCAT 1086

Qy         436 GluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAla 455
               |||||||||   |||::::::||||||:::||||||||||||||||||||||||||||||
Db        1087 GAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGCA 1146

Qy         456 SerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLys 475
               |||:::||||||||||||:::|||   :::||||||||||||||||||||||||||||||
Db        1147 TCATACTTGGAGAGGAAGAGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAAA 1206

Qy         476 AsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSerGln 494
               |||||||||||||||:::||||||||||||:::|||   :::|||||||||::::::|||
Db        1207 AATGAGAATTATAAAGCCCTCGTCAGAAGGAATTATGGGAAACGCAGAGGCAGAAACCAG 1266

Qy         495 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArg 514
               |||      :::      |||:::::::::   :::::::::::::::         :::
Db        1267 CAAATTGCTCGACCCTCGCAAGAAGAAAAAGTAGAAGAAAAAGAAGAG------GATAAA 1320

Qy         515 SerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGlu 534
               :::   :::   ||||||   |||:::||||||:::   |||   ||||||:::|||:::
Db        1321 GCAGAAAAAACAGAAAAAAAAGAAGAAGAAAAGAAAGATGAAGAGGAAAAAGATGAAAAA 1380

Qy         535 LysProGluValGluAsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSer 554
               :::         ||||||   ||||||         |||:::|||   |||   |||:::
Db        1381 GAAGACTCCAAAGAAAATACCAAGGAA---------AAGGACAAGATAGATGGTACAGCA 1431

Qy         555 GlyGluAspAsnAspGluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSer 574
                  |||:::   :::|||:::|||            :::|||||||||||||||||||||
Db        1432 ---GAAGAAACTGAGGAAAGAGAGCAAGCCACACCCCGGGGGCGAAAGACTGCCAACAGT 1488

Qy         575 GlnGlyArgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGlu 594
               ||||||||||||||||||||||||||||||||||||   |||||||||   :::
Db        1489 CAGGGCCGCCGTAAGGGCCGGATCACCAGGTCCATGACAAACGAAGCTGCAGCTGCCAGT 1548

Qy         595 AlaIleThrProGlnGlnSer---------------------------AlaGluLeu 604
               |||               :::                                |||
Db        1549 GCTGCAGCCGCAGCGGCTACTGAAGAGCCCCCACCACCTCTGCCACCGCCACCAGAACCC 1608

Qy         605 AlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLys 624
                  |||   |||      |||:::|||||||||||||||||||||||||||   ||||||
Db        1609 ATTTCTACAGAGCCTGTGGAGACCTCTCGATGGACAGAAGAAGAAATGGAAGTTGCTAAA 1668

Qy         625 LysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLys 644
               |||||||||:::||||||||||||||||||:::|||||||||:::|||||||||:::|||
Db        1669 AAAGGTCTAGTAGAACATGGTCGTAACTGGGCAGCAATTGCTAAAATGGTGGGAACGAAA 1728

Qy         645 ThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGlu 664
               :::   :::||||||||||||||||||||||||||||||:::|||   |||||||||
Db        1729 AGTGAAGCTCAATGTAAAAACTTCTATTTTAACTATAAAAGGCGACACAATCTTGACAAC 1788

Qy         665 IleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLys 684
               :::|||||||||||||||   |||      ::::::            |||
Db        1789 CTCTTACAGCAGCATAAACAGAAAACTTCACGAAAACCTCGTGAAGAGCGAGATGTGTCT 1848

Qy         685 LysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValValGluAspGluGlu 704
               :::   :::      ||||||      :::|||               |||||||||:::
Db        1849 CAATGTGAAAGTGTCGCTTCCACTGTTTCTGCT------------CAGGAGGATGAAGAT 1896

Qy         705 MetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeu 724
               :::|||||||||            ||||||||||||      |||::::::|||
Db        1897 ATTGAAGCCTCC------------AATGAAGAAGAAAATCCAGAAGACAGCGAAGGT--- 1941

Qy         725 HisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAlaThrValAsnAsn 744
                                                                        |||
Db        1942 ---------------------------------------------------GCAGAAAAT 1950

Qy         745 SerSerAspThrGluSerIleProSerProHis---ThrGluAlaAlaLysAspThrGly 763
               ||||||||||||||||||   |||||||||         ||||||   |||
Db        1951 AGTTCTGATACAGAAAGTGCTCCTTCTCCTTCACCAGTTGAAGCTGTCAAG--------- 2001

Qy         764 GlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProProGlyProPro 783
                                                               |||
Db        2002 ------------------------------------------------CCCAGCGAGGAC 2013

Qy         784 ThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThr 803
               :::|||         |||||||||         ||||||            |||
Db        2014 AGTCCTGAAAATGCTACTTCTCGAGGAAACACAGAACCTGCGGTTGAGCTTGAG------ 2067

Qy         804 GlyAlaProThrProProProAlaPro---ProSerProSer---AlaProProProVal 821
                     ||||||         ||||||      |||||||||   |||   |||
Db        2068 ------CCCACCACGGAAACTGCACCCAGTACATCTCCCTCCTTAGCAGTTCCAAGTACA 2121

Qy         822 ValProLysGluGluLysGluGluGluThrAlaAlaAlaProProVal------------ 837
                  |||   |||::::::      ||||||               :::
Db        2122 AAACCAGCTGAAGATGAAAGTGTGGAGACCCAGGTGAATGACAGCATCAGTGCTGAGACA 2181

Qy         838 ---------------------------------------GluGluGlyGluGluGlnLys 844
                                                      |||||||||
Db        2182 GCAGAGCAGATGGATGTAGATCAGCAGGAGCACAGTGCTGAAGAGGGTTCTGTTTGTGAT 2241

Qy         845 ProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSer 864
               ||||||   |||   :::         |||:::         |||      |||
Db        2242 CCCCCACCCGCTACCAAA------GCTGACTCTGTGGACGTTGAAGTGAGGGTGCCAGAA 2295

Qy         865 GluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGlu 884
                           :::   |||            |||   :::|||   :::   |||
Db        2296 AACCATGCATCTAAAGTTGAAGGTGATAATACCAAAGAAAGAGACTTGGATAGAGCC--- 2352

Qy         885 AlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThr 904
                                       :::||||||   |||
Db        2353 ------------------------AGTGAGAAGGTGGAACCTAGAGATGAAGATTTGGTG 2388

Qy         905 ThrAlaLysSerSerGlyAla------------ProGln---AspSerAspSerSerAla 919
                  |||:::         |||            ||||||   |||:::||||||||||||
Db        2389 GTAGCTCAGCAAATAAATGCCCAAAGGCCCGAGCCCCAGTCAGACAATGATTCCAGTGCC 2448

Qy         920 ThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSer 939
               ||||||||||||||||||      :::   :::|||      :::   :::   :::
Db        2449 ACGTGCAGCGCTGATGAG------GATGTGGATGGAGAGCCAGAGAGGCAGAGAATGTTT 2502

Qy         940 Pro---------ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSer 956
               |||         :::||||||||||||   |||||||||               :::|||
Db        2503 CCTATGGACTCAAAGCCTTCACTGTTAAACCCCACTGGATCTATACTCGTC---TCATCT 2559

Qy         957 ProGlnLys------ProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIlePro 974
               |||   |||      ||||||||||||   ||||||:::   |||||||||   ||||||
Db        2560 CCGTTAAAACCAAATCCACTGGATCTGCCACAGCTTCAGCATCGAGCTGCTGTTATCCCA 2619

Qy         975 ProIle------------------------------------------------------ 976
               |||:::
Db        2620 CCAATGGTATCCTGCACCCCATGTAACATACCAATTGGAACCCCAGTGAGCGGCTATGCT 2679

Qy         977 ------------GlnValThrLysValHisGluProProArgGluAspAlaAlaProThr 992
                              :::      :::||||||            :::
Db        2680 CTCTACCAGCGACACATTAAAGCAATGCATGAGTCAGCACTCCTGGAG------------ 2727

Qy         993 LysProAlaProProAlaProProProProGlnAsnLeuGlnProGluSerAspAlaPro 1012
                                             :::      |||   :::   |||
Db        2728 ------------------------GAGCAGCGGCAGAGACAAGAACAGATAGATTTGGAA 2763

Qy        1013 GlnGlnProGlySerSerProArgGlyLysSerArgSerProAlaProProAlaAspLys 1032
                  :::      :::||||||   |||   |||:::||||||            ::::::
Db        2764 TGTAGAAGTTCTACAAGTCCATGTGGCACATCCAAGAGTCCA------------AACAGA 2811

Qy        1033 GluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpThrSerGly 1052
               |||                                             |||
Db        2812 GAG---------------------------------------------TGG--------- 2817

Qy        1053 LeuProPheProValProProArgGluValIleLysAlaSerProHisAlaProAspPro 1072
                                       ||||||::::::   :::||||||
Db        2818 ------------------------GAAGTCCTTCAGCCTGCTCCACAT------------ 2841

Qy        1073 SerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAla 1092
                        :::      |||   |||      ||||||               |||
Db        2842 CAAGTGATAACTAATCTCCCTGAAGGCGTTCGGCTTCCG---------------ACAACT 2886

Qy        1093 ArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSerSerAla 1112
               ||||||      |||   ||||||   |||   :::            ::::::|||
Db        2887 CGACCAACCAGGCCACCGCCCCCTCTCATCCCGTCATCCAAAACCACAGTGGCTTCAGAA 2946

Qy        1113 LysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGly------------ 1128
               |||   ||||||   :::            |||:::||||||||||||
Db        2947 AAA---CCATCTTTTATA------ATGGGAGGCTCCATCTCACAGGGAACACCAGGCACT 2997

Qy        1129 ---MetSerValGlnLeuHisValProTyrSerGluHisAlaLysAlaPro---ValGly 1146
                  ::::::                  |||::::::            |||   ||||||
Db        2998 TATTTGACTTCTCATAATCAGGCTTCCTACACTCAAGAAACACCCAAGCCGTCAGTGGGA 3057

Qy        1147 ProValThrMetGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGly 1166
                  :::::::::|||||||||         :::      |||   |||
Db        3058 TCTATCTCTCTTGGACTGCCACGGCAACAGGAATCTGCCAAATCAGCTACTTTGCCCTAC 3117

Qy        1167 ValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyVal 1186
               :::|||||||||:::   |||||||||   |||         ||||||   |||   |||
Db        3118 ATCAAGCAGGAAGAATTTTCTCCCCGAAGCCAAAACTCACAACCTGAGGGTCTGTTGGTC 3177

Qy        1187 ProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGly 1206
                     ||||||         |||:::||||||||||||   |||::::::      |||
Db        3178 ---AGGGCCCAACATGAAGGTGTAGTCAGAGGTACCGCA---GGAGCCATACAAGAAGGA 3231

Qy        1207 SerIleThrLysGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArg 1226
               |||||||||:::|||   |||::::::::::::      :::|||      :::   |||
Db        3232 AGTATAACTCGGGGAACTCCAACCAGCAAAATTTCAGTGGAGAGCATTCCATCCCTACGG 3291

Qy        1227 GlySerIleThrHisGlyThrProAla---------------------------AspValLeu 1238
               ||||||||||||   ||||||||||||                        :::   |||
Db        3292 GGCTCTATCACTCAGGGCACCCCGGCTCTGCCCCAGACTGGCATACCAACAGAGGCTTTG 3351

Qy        1239 TyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGly 1258
                  ||||||:::|||:::|||:::      |||||||||            ::::::|||
Db        3352 GTGAAGGGGTCCATTTCGAGAATGCCCATTGAAGACAGCAGTCCT------GAGAAAGGC 3405

Qy        1259 ArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeu 1278
               ||||||::::::      |||||||||||||||||||||||||||   ||||||:::|||
Db        3406 AGAGAGGAAGCTGCATCCAAAGGCCATGTTATTTATGAAGGCAAAAGTGGACATATCTTG 3465

Qy        1279 SerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSer 1298
               ||||||:::         :::                     :::|||      |||
Db        3466 TCATATGATAATATTAAGAATGCC---------------CGAGAAGGGACTAGGAGTCCA 3510

Qy        1299 GlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVal 1318
                        ||||||   :::      ||||||:::|||:::   :::||||||   :::
Db        3511 AGAACAGCTCATGAAATCAGTTTA---AAGAGAAGCTATGAATCAGTGGAAGGAAATATA 3567

Qy        1319 GlyArgAlaIleSer------------------SerAlaSerIleGluGlyLeuMetGly 1332
                  :::   :::|||                  ||||||   :::|||||||||:::
Db        3568 AAGCAAGGGATGTCAATGAGGGAGTCTCCTGTATCAGCACCGTTAGAGGGGCTGATATGC 3627

Qy        1333 ArgAlaIleProProGluArgHisSerProHis---HisLeuLysGluGlnHisHisIle 1351
               ||||||:::|||      |||   |||||||||      |||||||||:::      :::
Db        3628 CGAGCATTACCC------AGGGGGAGTCCTCATTCTGACCTCAAAGAAAGGACTGTATTG 3681

Qy        1352 ArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeu 1371
                  |||||||||   ||||||   ||||||:::      |||:::   ||||||   |||
Db        3682 TCTGGCTCCATAATGCAGGGGACACCAAGAGCAACAACTGAAAGCTTTGAAGATGGCCTT 3741

Qy        1372 ArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSerArg 1391
                  :::      |||   :::|||||||||                  ||||||   |||
Db        3742 ---AAATATCCCAAACAAATTAAAAGGGAAAGT---------------CCTCCCATACGA 3783

Qy        1392 AspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHis 1411
                           |||                           |||   ||||||   :::
Db        3784 GCATTTGAAGGTGCCATT---------------------ACCAAAGGAAAACCA---TAT 3819

Qy        1412 GluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGlu 1431
               :::|||   :::   |||:::||||||   |||||||||||||||||||||||||||:::
Db        3820 GATGGC---ATCACCACCATCAAAGAAATGGGGCGTTCCATTCATGAGATTCCAAGGCAA 3876

Qy        1432 GluLeu---------------ArgHisThrProGluLeuProLeuAlaProArgProLeu 1446
               ::::::               |||   |||||||||:::      :::   ||||||:::
Db        3877 GATATTTTAACTCAGGAAAGTCGGAAAACTCCAGAAGTGGTCCAGAGCACACGGCCGATA 3936
```

```
Qy      1447 LysGluGlySerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThr 1466
                ||||||||||||:::||||||||||||:::|||:::|||         |||
Db      3937 ATTGAGGGTTCCATTTCCCAGGGCACACCAATAAAGTTTGACAACAAC---TCAGGTCAA 3993

Qy      1467 GlySerLysLysHisAspValArgSerLeuIleGlySerProGlyArgThrPheProPro 1486
                :::   ||||||:::|||:::|||||||||      |||   :::
Db      3994 TCTGCCATCAAACACAATGTCAAATCCTTAATCACGGGGCCTAGCAAACTATCCCGTGGA 4053

Qy      1487 ValHisProLeuAspValMetAlaAsp---AlaArgAlaLeuGluArgAlaCysTyrGlu 1505
             :::   ||||||:::::::::   :::      :::   :::||||||      ||||||
Db      4054 ATGCCTCCGCTGGAAATTGTGCCAGAGAACATAAAAGTGGTAGAACGGGGAAAATATGAG 4113

Qy      1506 ---------------GluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
                            |||:::::::::||||||         :::         |||||||||
Db      4114 GATGTGAAAGCAGGCGAGACCGTGCGTTCCCGGCACACGTCAGTGGTAAGCTCTGGCCCC 4173

Qy      1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
             |||:::   |||                  :::   |||      |||   :::   ||||||
Db      4174 TCCGTTCTTAGGTCCACA---------CTGCATGAAGCTCCCAAAGCACAACTGAGCCCT 4224

Qy      1541 LeuThrTyrGluAspHisGlyAla---------------ProPheAlaGlyHisLeuPro 1555
                   |||:::|||      |||                  :::         :::
Db      4225 GGGATTTATGATGACACCAGTGCACGGAGGACCCCTGTGAGTTATCAAAACACCATGTCC 4284

Qy      1556 ArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSer 1575
             ||||||||||||:::      |||                              :::|||
Db      4285 AGAGGCTCACCCATGATGAACAGA---------------------ACTTCTGATGTTTCT 4323

Qy      1576 SerSerLysAla---SerGlnAspArgLysLeuThrSerThrProArgGlu--------- 1591
             |||:::|||:::   :::   :::||||||   |||   ||||||   :::
Db      4324 TCTAACAAGTCTACCAATCATGAAAGGAAATCGACACTGACCCCTACCCAGAGGGAAAGT 4383

Qy      1592 ---IleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPro 1610
                   ||||||||||||         |||               |||      ||||||
Db      4384 ATCCCAGCGAAGTCTCCAGTGCCTGGGGTGGACCCTGTCGTGAGCCAC------AGTCCG 4437

Qy      1611 TyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeu 1630
             ::::::         ||||||   :::         ::::::||||||||||||:::|||
Db      4438 TTTGATCCCCATCACAGAGGCAGCACTGCAGGCGAGGTTTATCGGAGCCACCTGCCCACG 4497

Qy      1631 AlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyr 1650
                   ||||||   ::::::|||            ||||||   ||||||||||||
Db      4498 CACTTGGATCCA---GCCATGCCTTTTCACAGGGCTTTGGATCCTGCGGCTGCTTACCTG 4554

Qy      1651 LeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIle 1670
                   |||   |||:::|||   |||   ||||||      |||      |||
Db      4555 TTTCAGAGACAGCTTTCACCAACTCCAGGTTACCCAAGTCAGTATCAGCTTTAC------ 4608

Qy      1671 ArgGlyTyrProAspThrAlaAlaLeuGluAsn---ArgGlnThrIleIleAsnAspTyr 1689
                                  |||:::||||||   ||||||||||||:::|||||||||
Db      4609 ---------------------GCAATGGAGAACACAAGACAGACAATCTTAAATGATTAC 4647

Qy      1690 IleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMet 1709
             ||||||||||||||||||      |||                     |||   |||:::
Db      4648 ATTACCTCACAACAGATGCAAGTGAACTTG------------------CGTCCAGATGTG 4689

Qy      1710 LeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArg 1729
                |||||||||||||||||||||      |||   |||   |||   |||      |||
Db      4690 GCCAGAGGACTCTCCCCAAGAGAGCAGCCACTGGGTCTCCCATACCCAGCA---ACGAGA 4746

Qy      1730 GlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGly 1749
             |||||||||||||||:::   :::|||         :::|||||||||         |||
Db      4747 GGAATCATTGACCTGACCAATATGCCTCCA---ACAATTTTAGTGCCTCATCCAGGGGGA 4803

Qy      1750 ThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSer 1769
             |||            |||||||||:::   |||:::|||               |||
Db      4804 ACAAGCACTCCTCCCATGGACAGAATCACTTATATTCCTGGTACACAGATTACTTTCCCT 4863

Qy      1770 SerArg---HisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysPro 1788
                |||   ::::::|||:::   :::|||||||||   ||||||||||||
Db      4864 CCCAGGCCGTACAACTCTGCTTCCATGTCTCCAGGACACCCAACACACCTT--------- 4914

Qy      1789 ThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArg 1808
                      :::|||:::||||||||||||:::|||:::||||||::::::|||:::|||
Db      4915 GCAGCTGCTGCAAGTGCTGAGAGGGAACGGGAACGGGAGCGGGAGAAGGAGCGGGAGCGG 4974

Qy      1809 GluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArg 1828
             ||||||         |||      :::::::::   :::                  |||
Db      4975 GAACGG---------ATTGCTGCAGCTTCCTCCGACCTC------------TACCTGCGG 5013

Qy      1829 ProGlyThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSer 1848
             ||||||:::||||||
Db      5014 CCAGGCTCAGAACAG---------------------------------------CCT 5031

Qy      1849 SerArgProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGln 1868
                ||||||   ||||||   :::            ||||||   |||   |||||||||
Db      5032 GGCCGACCTGGCAGTCATGGATATGTTCGCTCCCCTTCCCCT---TCAGTAAGAACTCAG 5088

Qy      1869 AspAla---LeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIle 1887
             :::      |||||||||||||||||||||         |||            :::|||
Db      5089 GAGACCATGTTGCAACAGAGACCCAGTGTTTTCCAAGGAACCAATGGAACCAGTGTAATC 5148

Qy      1888 ThrAlaValGluProSerLysProThrValLeuArgSerThrSerThrSerSerProVal 1907
             |||   ::::::|||:::            :::                     |||
Db      5149 ACACCTTTGGATCCAACTGCTCAGCTACGAATCATGCCACTGCCTGCTGGGGGCCCTTCA 5208

Qy      1908 ------------ArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGly 1923
                            ||||||:::   :::      |||            |||
Db      5209 ATAAGCCAAGGCCTGCCAGCCTCCCGTTACAACACTGCTGCGGATGCC---CTGGCTGCT 5265

Qy      1924 ThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys--------- 1940
                :::|||                           |||         |||
Db      5266 CTTGTGGATGCTGCAGCTTCTGCACCCCAGATGGATGTGTCCAAAACAAAAGAGAGTAAG 5325

Qy      1941 ---GluAlaProArgValAlaArgProGluArgProArgAlaAsp--------------- 1954
                ||||||   |||:::            |||   |||:::
Db      5326 CATGAAGCTGCCAGGTTAGAAGAAAATTTGAGAAGCAGGTCAGCAGCAGTTAGTGAACAG 5385

Qy      1955 ---------------------------------------------------------Thr 1955
                                                                      |||
Db      5386 CAGCAGCTAGAGCAGAAAACCCTGGAGGTGGAGAAGAGATCTGTTCAGTGTTTATACACT 5445

Qy      1956 GlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSer--- 1974
                   ||||||   :::      |||               :::|||   ||||||
Db      5446 TCTTCAGCCTTTCCAAGTGGCAAGCCC---------------CAGCCTCATTCTTCAGTA 5490

Qy      1975 ------ProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAla 1992
                            |||   :::      |||      ||||||   |||   :::
Db      5491 GTTTATTCTGAGGCTGGGAAAGATAAAGGGCCT------CCTCCAAAATCCAGATATGAG 5544

Qy      1993 ThrIleAlaArgThrProAlaLysAsn------LeuAlaProHisHisAlaSerProAsp 2010
                      ||||||      |||            |||   :::
Db      5545 GAAGAGCTAAGGACCAGAGGGAAGACTACCATTACTGCAGCTAACTTCATAGACGTGATC 5604

Qy      2011 ProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysPro 2030
                            ||||||      |||               :::||||||
Db      5605 ATCACCCGGCAAATTGCCTCGGACAAGGATGCGAGGGAACGTGGCTCTCAAAGTTCAGAC 5664

Qy      2031 PheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerPro 2050
                |||                     ||||||      |||
Db      5665 TCTTCT------------------AGTAGCTTATCTTCTCACAGGTATGAAACACCTAGC 5706

Qy      2051 GluGlyValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeu 2070
             :::   :::|||   :::||||||   |||||||||:::         :::|||
Db      5707 GATGCTATTGAGGTGATAAGTCCTGCCAGCTCACCTGCGCCACCCCAGGAGAAACTGCAG 5766

Qy      2071 ProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGln 2090
                         |||   :::      :::      |||   :::         :::
Db      5767 ACCTATCAGCCAGAGGTTGTTAAGGCAAATCAAGCGGAAAATGATCCTACCAGACAATAT 5826

Qy      2091 ProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuPro 2110
                ||||||                           |||   |||   ||||||
Db      5827 GAAGGACCA---------------------------TTACATCACTATCGACCA------ 5853

Qy      2111 GluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGly--------------- 2125
                   |||   |||   ||||||         |||      |||
Db          5854 ---CAGCAGGAATCACCATCTCCCCAACAACAGCTGCCCCCTTCTTCACAGGCAGAGGGA 5910

Qy          2126 ---------ValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluVal 2142
                          |||            |||::::::|||||||||   ||||||   ::::::
Db          5911 ATGGGGCAAGTGCCCAGGACCCATCGGCTGATCACACTTGCTGATCACATCTGTCAAATT 5970

Qy          2143 IleThrGlnAspTyrThrArgHisHis---------------ProGlnGlnLeuSerAla 2157
                 ||||||||||||:::   |||:::                  |||||||||
Db          5971 ATCACACAAGATTTTGCTAGAAATCAAGTTTCCTCGCAGACTCCCCAGCAG--------- 6021

Qy          2158 ProLeuProAlaProLeuTyrSerPheProGlyAla------SerCysProValLeuAsp 2175
                 |||      :::         :::   |||   |||      |||   ||||||
Db          6022 CCTCCTACTTCTACATTCCAGAACTCACCTTCTGCTTTGGTATCTACACCTGTG------ 6075

Qy          2176 LeuArgArgProProSerAspLeuTyrLeuProProProAsp---------------His 2190
                    |||         |||:::   |||   |||                        |||
Db          6076 ---AGGACTAAAACATCAAACCGTTACAGCCCAGAATCCCAGGCTCAGTCTGTCCATCAT 6132

Qy          2191 GlyAlaPro---AlaArgGlySerProHis---------------SerGluGlyGlyLys 2204
                       |||   :::|||   ||||||                  |||   |||   :::
Db          6133 CAAAGACCAGGTTCAAGGGTCTCTCCAGAAAATCTTGTGGACAAATCCAGGGGAAGTAGG 6192

Qy          2205 ArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluPro 2224
                             |||   ::::::   |||               :::      ||||||
Db          6193 CCTGGAAAATCCCCAGAGAGGAGTCACGTC---------TCTTCGGAGCCCTACGAGCCC 6243

Qy          2225 ValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeu 2244
                 :::|||||||||:::                  |||   :::               |||
Db          6244 ATCTCCCCACCCCAG------GTTCCGGTTGTGCATGAGAAACAGGACAGCTTGCTGCTC 6297

Qy          2245 LeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlyThrLysSerProGlyAsn 2264
                 |||   :::   |||   :::            |||      ::::::|||||||||:::
Db          6298 TTGTCTCAGAGGGGCGCAGAGCCTGCAGAGCAGAGGAATGATGCCCGCTCACCAGGGAGT 6357

Qy          2265 ThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLys 2284
                    |||      |||:::||||||:::||||||   |||:::   |||   |||||||||
Db          6358 ATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAATACATCACCCATGGTTAAA 6414

Qy          2285 SerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyr 2304
                 ||||||||||||||||||   :::|||||||||:::            :::   :::
Db          6415 TCAAAGAAGCAGGAGATTTTTCGTAAGTTGAACTCCTCTGGTGGAGGTGACTCTGATATG 6474

Qy          2305 AsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeu 2324
                       :::||||||||||||||||||||||||:::||||||:::|||   :::|||
Db          6475 GCAGCTGCTCAGCCAGGAACTGAGATCTTTAATCTGCCAGCAGTTACTACGTCAGGCTCA 6534

Qy          2325 MetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAla 2344
                 ::::::   |||      :::      :::   ||||||   |||:::|||||||||
Db          6535 GTTAGCTCTAGAGGCCATTCTTTTGCTGATCCTGCCAGT---AATCTTGGGCTGGAAGAC 6591

Qy          2345 IleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu--------------- 2359
                 ||||||||||||||||||||||||   :::|||         |||
Db          6592 ATTATCAGGAAGGCTCTCATGGGAAGCTTTGATGACAAAGTTGAGGATCATGGAGTTGTC 6651

Qy          2360 ---SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuPro 2378
                    |||   |||:::                  |||      |||:::   |||:::
Db          6652 ATGTCCCAGCCTATGGGAGTA---------GTGCCTGGTACTGCCAACACCTCAGTT--- 6699

Qy          2379 AlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGly 2398
                             :::|||:::   :::   |||   :::            ||||||
Db          6700 ------------GTGACCAGTGGTGAGACACGAAGAGAGGAAGGGGACCCATCACCTCAT 6747

Qy          2399 GlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerPro 2417
                    |||      |||   |||:::      :::   :::|||||||||:::|||||||||
Db          6748 TCAGGAGTTTGCAAACCAAAGCTGATCAGCAAGTCAAACAGCAGGAAATCTAAGTCTCCT 6807

Qy          2418 AlaPro------GlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSer 2435
                    |||      |||            :::||||||   |||||||||||||||||||||
Db          6808 ATACCTGGGCAAGGCTACTTAGGAACGGAACGGCCCTCTTCAGTCTCCTCTGTACATTCA 6867

Qy          2436 GluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSer 2455
                 |||||||||   :::|||:::||||||               ||||||||||||||||||
Db          6868 GAAGGGGATTACCATAGGCAGACGCCA------GGGTGGCCTGGGAAGACAGGCCCTCT 6921

Qy          2456 SerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGlnAlaGlyVal 2475
                 |||   |||||||||   ||||||||||||||||||   ||||||:::         :::
Db          6922 TCAACAGGCTCAACTCAGTTTCCTTATAACCCTCTGACTATGCGGATG---------CTC 6972

Qy          2476 MetAlaSerProProProProGlyLeuProAlaGlySerGly---ProLeuAlaGlyPro 2494
                    ::::::||||||   |||         |||   |||            |||   |||
Db          6973 AGCAGTACTCCACCAACACCGATTGCATGTGCTCCCTCTGCGGTGAACCAAGCAGCTCCT 7032

Qy          2495 His---------HisAlaTrpAspGluGluProLysProLeuLeuCysSerGlnTyrGlu 2511
                 |||               |||:::   ||||||   |||||||||   :::|||||||||
Db          7033 CACCAACAGAACAGGATCTGGGAGCGAGAGCCTGCCCCACTGCTCTCAGCACAGTACGAG 7092

Qy          2512 ThrLeuSerAspSerGlu 2517
                 |||||||||||||||:::
Db          7093 ACCCTGTCGGATAGTGAT 7110

Search completed: November 11, 2005, 20:52:48
Job time : 36545 secs
```

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2005  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:          November 11, 2005, 09:34:57 ; Search time 24239 Seconds
                                          (without alignments)
                                          3952.629 Million cell updates/sec

Title:           US-09-522-753-5
Perfect score:   13215
Sequence:        1 MSGSTQLVAQTWRATEPRYP..........WDEEPKPLLCSQYETLSDSE 2517

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters:      68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09522753/runat_02112005_173640_20807/app_query.fasta_1.2695
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09522753_@CGN_1_1_10059_@runat_02112005_173640_20807 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  gb_est1:*
                 2:  gb_est2:*
                 3:  gb_htc:*
                 4:  gb_est3:*
                 5:  gb_est4:*
                 6:  gb_est5:*
                 7:  gb_est6:*
                 8:  gb_gss1:*
                 9:  gb_gss2:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                   Description
---------------------------------------------------------------------------
    1     8226    62.2   7372  9  AY412686             AY412686 Homo sapi
    2   7109.5    53.8   7013  9  AY412688             AY412688 Mus muscu
    3     6533    49.4   6836  9  AY412687             AY412687 Pan trogl
    4     2566    19.4   2000  3  BC033087             BC033087 Homo sapi
    5     2264    17.1   1362  3  BC020427             BC020427 Homo sapi
    6   1643.5    12.4   1966  3  BC026028             BC026028 Homo sapi
    7     1634    12.4   1850  3  BC035748             BC035748 Homo sapi
    8   1626.5    12.3   1835  3  BC029627             BC029627 Mus muscu
    9   1624.5    12.3   1819  3  BC062162             BC062162 Mus muscu
```

```
   10     1624    12.3   1093  5  BM909096             BM909096 AGENCOURT
   11     1603    12.1   1996  3  BC026623             BC026623 Mus muscu
   12     1559    11.8    965  4  BM555371             BM555371 AGENCOURT
   13     1540    11.7   1126  4  BM553310             BM553310 AGENCOURT
   14   1480.5    11.2   1057  4  BM477568             BM477568 AGENCOURT
   15   1479.5    11.2   1051  4  BM423558             BM423558 AGENCOURT
   16   1471.5    11.1    937  5  BQ949384             BQ949384 AGENCOURT
   17     1448    11.0    846  5  BU557144             BU557144 AGENCOURT
   18   1442.5    10.9    984  5  BQ690869             BQ690869 AGENCOURT
   19   1440.5    10.9   1159  4  BM802749             BM802749 AGENCOURT
   20     1420    10.7   1022  5  BM910785             BM910785 AGENCOURT
   21   1382.5    10.5   1030  4  BM558844             BM558844 AGENCOURT
   22   1378.5    10.4   1010  5  BM915686             BM915686 AGENCOURT
   23     1374    10.4    887  5  BU542258             BU542258 AGENCOURT
   24     1368    10.4    875  7  CK769773             CK769773 957873 MA
   25     1368    10.4   1067  4  BM472005             BM472005 AGENCOURT
   26   1366.5    10.3    868  5  BX368971             BX368971 BX368971
   27   1348.5    10.2    923  5  BU184403             BU184403 AGENCOURT
   28   1346.5    10.2   1038  5  BM910704             BM910704 AGENCOURT
   29     1338    10.1    887  6  CA979881             CA979881 AGENCOURT
   30     1334    10.1    923  5  BX390462             BX390462 BX390462
   31     1332    10.1    875  5  BQ711119             BQ711119 AGENCOURT
   32   1331.5    10.1    842  5  BU224569             BU224569 603400211
   33     1326    10.0   1094  4  BM560912             BM560912 AGENCOURT
   34   1325.5    10.0    905  5  BQ897825             BQ897825 AGENCOURT
   35     1321    10.0    882  5  BU180236             BU180236 AGENCOURT
   36   1305.5     9.9    898  5  BU172348             BU172348 AGENCOURT
   37     1305     9.9   1045  4  BM560255             BM560255 AGENCOURT
   38   1297.5     9.8    833  7  CK778015             CK778015 965166 MA
   39     1274     9.6    875  5  BQ691710             BQ691710 AGENCOURT
   40     1272     9.6   3575  3  HSM802410            AL137641 Homo sapi
   41     1267     9.6    769  5  BX368972             BX368972 BX368972
   42   1262.5     9.6    915  7  CF272442             CF272442 AGENCOURT
   43     1260     9.5   1075  4  BM461469             BM461469 AGENCOURT
   44   1259.5     9.5    785  7  CO896175             CO896175 BovGen_24
   45   1256.5     9.5    928  5  BU164114             BU164114 AGENCOURT


                                  ALIGNMENTS


RESULT 1
AY412686
LOCUS       AY412686              7372 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Homo sapiens NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412686
VERSION     AY412686.1  GI:39768651
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 7372)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 7372)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .7372
```

```
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
     gene           <1. .>7372
                    /gene="NCOR2"
                    /locus_tag="HCM4640"
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         7372
Score:                  8226.00         Matches:        1708
Percent Similarity:     69.56%          Conservative:   1
Best Local Similarity:  69.52%          Mismatches:     748
Query Match:            62.25%          Indels:         0
DB:                     9               Gaps:           0

US-09-522-753-5 (1-2517) x AY412686 (1-7372)

Qy            1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db            1 ATGTCGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 60

Qy           21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 120

Qy           41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 180

Qy           61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGNNNNNNN 240

Qy           81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100

Db          241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300

Qy          101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120

Db          301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360

Qy          121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
                                                                      ||||||
Db          361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACCGTAGC 420

Qy          141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 480

Qy          161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 540

Qy          181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 600

Qy          201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 660

Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                |||||||||||||||||||||||||||||||||||||||||||||
Db          661 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGNNNNNNNNNNNNNNN 720

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                                                          ||||||||||||||||||
Db          721 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGCTGTACAACCAGCCC 780

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 840

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          841 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGAGCAGAAGTTCTGCCAG 900

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db          901 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 960

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          961 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1020

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1021 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1080

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1081 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1140

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1141 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1200

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1201 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1260

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1261 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1320

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||
Db         1321 CGGAGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1380

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1381 NNNACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1440

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||
Db         1441 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGNNNNNNNNNNNNNNNNNN 1500

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520

Db         1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540

Db         1561 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1620

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                                     |||||||||||||||||||||||||||||||||||||||
Db         1621 NNNNNNNNNNNNNNNNNNNNGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAG 1680

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1681 AAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGC 1740

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1741 CGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAG 1800

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1801 AGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATG 1860

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1861 GAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATG 1920
```

```
Qy        641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1980

Qy        661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCGCGG 2040

Qy        681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 AGGAAGAAGAAGAAAGCGCCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTG 2100

Qy        701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAG 2160

Qy        721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
              ||||||
Db       2161 GCTGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2220

Qy        741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760

Db       2221 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2280

Qy        761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780

Db       2281 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2340

Qy        781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800

Db       2341 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2400

Qy        801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820

Db       2401 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2460

Qy        821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
                                    ||||||||||||||||||||||||||||||||||||||
Db       2461 NNNNNNNNNNNNNNNNNNNNNNAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGG 2520

Qy        841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAG 2580

Qy        861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 CCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCG 2640

Qy        881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySer 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGC 2700

Qy        901 GlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThr 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCTACC 2760

Qy        921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCA 2820

Qy        941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLysPro 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCA 2880

Qy        961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnValThrLys 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAA 2940

Qy        981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCG 3000

Qy       1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGG 3060

Qy       1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
              |||||||||||||||||||||||||||||||||||||||
Db       3061 GGCAAGAGCAGGAGCCCGGCACCCCCGCCGACAAGGAGGNNNNNNNNNNNNNNNNNNNNN 3120

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060

Db       3121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3180

Qy       1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080

Db       3181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3240

Qy       1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 NGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 3300

Qy       1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 3360

Qy       1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 3420

Qy       1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 3480

Qy       1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 3540

Qy       1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 CCGGAGAGCCTGGGGGTGCCCACAGCCCAGGAGGCGTCCGTGCTGAGAGGGACAGCTCTG 3600

Qy       1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GGCTCAGTTCCGGGCGGAAGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGAC 3660

Qy       1221 SerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrLys 1240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 AGCGCCATCACATACCGCGGCTCCATCACCCACGGCACGCCAGCTGACGTCCTGTACAAG 3720

Qy       1241 GlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GGCACCATCACCAGGATCATCGGCGAGGACAGCCCGAGTCGCTTGGACCGCGGCCGGGAG 3780

Qy       1261 AspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyr 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 GACAGCCTGCCCAAGGGCCACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTAT 3840

Qy       1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro 1300
              |||
Db       3841 GAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3900

Qy       1301 ProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArg 1320

Db       3901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3960

Qy       1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
                                         |||||||||||||||||||||||||||||||||
Db       3961 NNNNNNNNNNNNNNNNNNNNNNNNNGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACAC 4020

Qy       1341 SerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGlyIlePro 1360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AGCCCCCACCACCTCAAAGAGCAGCACCACATCCGCGGGTCCATCACACAAGNNNNNNNN 4080

Qy       1361 ArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArg 1380
```

```
Db       4081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4140

Qy       1381 GluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLysThrGln 1400

Db       4141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4200

Qy       1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420

Db       4201 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4260

Qy       1421 AlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGluLeuPro 1440

Db       4261 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4320

Qy       1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460

Db       4321 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4380

Qy       1461 ThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGlySerPro 1480
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4381 NNNNNNNCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTCATCGGCAGCCCC 4440

Qy       1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4441 GGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAA 4500

Qy       1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4501 CGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGCAGCTCGGGGGGC 4560

Qy       1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4561 TCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCCGCGGCAGAGCCCC 4620

Qy       1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySerProVal 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4621 CTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGAGGTTCGCCCGTG 4680

Qy       1561 ThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSer 1580
              |||   |||||||||||||||||||||||||||
Db       4681 ACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGNNNNNNNNNNNNNNNNNNNNNNNNNN 4740

Qy       1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600

Db       4741 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4800

Qy       1601 ProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyValSerGly 1620

Db       4801 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4860

Qy       1621 ValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIleProArgGly 1640

Db       4861 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4920

Qy       1641 IleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
                                ||||||||||||||||||||||||||||||||||||||||||
Db       4921 NNNNNNNNNNNNNNNNCCGCTGCCTACTACCTGCCCCGACACCTGGCCCCCAACCCCACC 4980

Qy       1661 TyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4981 TACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACGGCGGCGCTGGAG 5040

Qy       1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAla 1700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db       5041 AACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGCC 5100

Qy       1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeu 1720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5101 ACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTG 5160

Qy       1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5161 GCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCCCAAGTGCCACACCTG 5220

Qy       1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5221 CCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATGGACCGCCTTGCCTAC 5280

Qy       1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5281 CTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCCCCACTCTCCCCAGGA 5340

Qy       1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5341 GGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAGCGGGAGCGAGACCGG 5400

Qy       1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5401 GATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTCACGTCCACCACGACG 5460

Qy       1821 ValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5461 GTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGC 5520

Qy       1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5521 AGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCCCATGCCCACCAGCAC 5580

Qy       1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5581 TCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCCAGTGTGCTTCACAAC 5640

Qy       1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       5641 ACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCCACGGTCCTGAGGTCC 5700

Qy       1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5701 ACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCCA 5760

Qy       1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5761 CTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAG 5820

Qy       1941 GluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5821 GAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTC 5880

Qy       1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5881 GCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAG 5940

Qy       1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5941 CCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAG 6000

Qy       2001 AsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAsp 2020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6001 AACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGAC 6060

Qy       2021 ProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6061 CCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCT 6120

Qy       2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
              |||
Db       6121 CTGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6180

Qy       2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSer 2080

Db       6181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6240

Qy       2081 HisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
                                                  ||||||||||||||||||||||||
Db       6241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGCCCCGTGAAGCTTGGCGGGGAGGCC 6300
```

```
Qy       2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6301 GCCCACCTCCCACACCTGCGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTC 6360

Qy       2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisIleSer 2140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6361 CAGACCGCCCCAGGGGTCAAAGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGT 6420

Qy       2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160

Db       6421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6480

Qy       2161 AlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArgProPro 2180

Db       6481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6540

Qy       2181 SerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSer 2200

Db       6541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6600

Qy       2201 GluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAsp 2220
                             |||||||||||||||||||||||||||||||||||||||||||||
Db       6601 NNNNNNNNNNNNNNGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGAC 6660

Qy       2221 GlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAla 2240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6661 GGTATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCT 6720

Qy       2241 ValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLys 2260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6721 GTGTACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAG 6780

Qy       2261 SerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSer 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6781 TCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCC 6840

Qy       2281 AlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsn 2300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6841 GCCATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAAT 6900

Qy       2301 GluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThr 2320
              ||||||||||||
Db       6901 GAGCCTGAATACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6960

Qy       2321 GlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMet 2340

Db       6961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 7020

Qy       2341 GlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSer 2360

Db       7021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 7080

Qy       2361 ProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAla 2380

Db       7081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 7140

Qy       2381 MetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGly 2400
                                                                    ||||||
Db       7141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGGCGGC 7200

Qy       2401 GlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGly 2420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7201 GGGAAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGC 7260

Qy       2421 LeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsn 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7261 CTGGCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAAC 7320

Qy       2441 ArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAla 2457
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7321 CGCCGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCA 7371

RESULT 2
```

```
AY412688
LOCUS       AY412688                7013 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Mus musculus NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412688
VERSION     AY412688.1  GI:39768653
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 7013)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 7013)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1..7013
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>7013
                     /gene="NCOR2"
                     /locus_tag="HCM4640"
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         7013
Score:                  7109.50         Matches:        1509
Percent Similarity:     63.77%          Conservative:   59
Best Local Similarity:  61.37%          Mismatches:     768
Query Match:            53.80%          Indels:         125
DB:                     9               Gaps:           17

US-09-522-753-5 (1-2517) x AY412688 (1-7013)

Qy            1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
                ||||||||||||||||||   |||||||||||||||||||||   |||||||||||||||
Db            1 ATGTCAGGATCCACACAGCCTGTGGCACAGACATGGCGGGCTGCTGAGCCCCGCTACCCA 60

Qy           21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
                ||||||   :::||||||||||||||||||||||||:::|||||||||||||||||||||
Db           61 CCCCATGGCATCTCCTACCCGGTGCAGATAGCCCGGTCCCACACGGACGTGGGGCTGCTT 120

Qy           41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
                |||||||||||||||   |||||||||   ||||||||||||||||||||||||||||||
Db          121 GAGTACCAACACCACCCCCGTGACTACACCTCACACCTGTCACCCGGTTCCATCATCCAG 180

Qy           61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
                |||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db          181 CCACAGAGGAGGCGGCCCTCACTGCTGTCAGAGTTCCAGCCTGGGAGTGAACGNNNNNNN 240

Qy           81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100

Db          241 NNNNNNNNNNNNNNNNN-NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 299

Qy          101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120

Db          300 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 359
```

```
Qy          121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
                                                                      ||||||
Db          360 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACCGTAGC 419

Qy          141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
                |||   |||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db          420 CTGGCAGGCAAGCTGGAGCCTGTGTCACCTCCCAGTCCCCCGCACGCTGACCCTGAGCTA 479

Qy          161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
                ||||||   |||   |||||||||||||||||||||||||||||||||||||||||||||
Db          480 GAGCTGGCGCCATCTCGACTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 539

Qy          181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          540 CGTGAGATCACCATGGTAGAGCAGCAGATCTCCAAGCTGAAGAAGAAGCAGCAACAGTTG 599

Qy          201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          600 GAGGAGGAGGCCGCCAAGCCGCCCGAACCCGAGAAGCCTGTGTCGCCACCACCCATAGAA 659

Qy          221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
                |||||||||||||||||||||||||||||||||||||||||||||
Db          660 TCAAAGCACCGAAGCCTGGTCCAGATCATCTACGATGAGAACCGGNNNNNNNNNNNNNNN 719

Qy          241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
                                                          ||||||||||||||||||
Db          720 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCTCTGTACAACCAGCCG 779

Qy          261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          780 TCTGACACACGCCAGTACCATGAAAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTG 839

Qy          281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
                ||||||||||||||||||||||||||||||||||||||||||:::|||:::|||||||||
Db          840 ATCTTGTACTTTAAGCGGAGGAACCACGCGCGCAAGCAGTGGGAACAGCGCTTCTGCCAG 899

Qy          301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
                ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db          900 CGCTATGACCAGCTCATGGAGGCGTGGGAGAAGAAGGTAGAGCGCATAGAGAACAATCCG 959

Qy          321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          960 CGAAGGAGGGCCAAGGAGAGCAAGGTGAGGGAGTACTACGAGAAACAGTTCCCGGAGATC 1019

Qy          341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1020 CGCAAGCAGCGGGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGTGGCAGTGGG 1079

Qy          361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1080 CTCTCCATGTCGGCTGCCCGCAGTGAGCATGAGGTTTCTGAGATCATTGATGGCTTGTCT 1139

Qy          381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1140 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGTTGTAC 1199

Qy          401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
                ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db         1200 GACGCGGACCAGCAGAGGATCAAGTTCATCAACATGAATGGACTCATGGATGACCCCATG 1259

Qy          421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
                ||||||||||||||||||||||||   |||||||||||||||||||||::::::||||||
Db         1260 AAGGTCTACAAGGACCGTCAGGTTACCAACATGTGGAGCGAGCAGGAGAGGGACACCTTC 1319

Qy          441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
                ||||||
Db         1320 CGTGAGAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1379

Qy          461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
                   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1380 NNNACGGTCGCTGAGTGTGTCCTCTATTACTACCTGACCAAGAAGAATGAAAATTACAAG 1439

Qy          481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
                ||||||||||||||||||||||||||||||||||||||||||
Db         1440 AGCTTGGTGAGGCGGAGCTATCGGCGCCGTGGCAAGAGCCAGNNNNNNNNNNNNNNNNNN 1499

Qy          501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520

Db         1500 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1559

Qy          521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540

Db         1560 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1619

Qy          541 AspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGlu 560
                                     |||||||||||||||||||||||||||||||||||||||
Db         1620 NNNNNNNNNNNNNNNNNNNNGGAGAAGACAGACGACACTTCTGGCGAGGACAACGATGAG 1679

Qy          561 LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGly 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1680 AAAGAGGCCGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAAGGCCGCCGCAAAGGC 1739

Qy          581 ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGln 600
                |||||||||||||||||||||||||||||||||   ||||||      ||||||||||||
Db         1740 CGTATCACGCGCTCCATGGCCAACGAGGCCAACCATGAGGAGACAGCCACCCCACAGCAA 1799

Qy          601 SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGluMet 620
                |||:::||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db         1800 AGTTCAGAGCTGGCTTCCATGGAGATGAACGAGAGTTCTCGCTGGACTGAGGAAGAGATG 1859

Qy          621 GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1860 GAGACAGCAAAGAAAGGCCTCCTGGAACATGGGAGGAACTGGTCAGCCATTGCCCGCATG 1919

Qy          641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1920 GTGGGCTCCAAGACCGTGTCCCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAG 1979

Qy          661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1980 AACCTGGACGAAATCCTTCAGCAGCACAAGCTAAAGATGGAGAAGGAGAGGAACGCTCGG 2039

Qy          681 ArgLysLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
                |||||||||||||||   |||||||||||||||||||||   ||||||||||||
Db         2040 AGGAAGAAGAAGAAGACCCCAGCTGCGGCGAGCGAGGAGACAGCCTTCCCACCTGCCGCT 2099

Qy          701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
                |||||||||||||||||||||||||||   |||   ||||||||||||:::   ||||||
Db         2100 GAGGACGAAGAGATGGAAGCATCAGGCGCAAGTGCCAATGAGGAAGAGCTGGCGGAGGAG 2159

Qy          721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
                ||||||
Db         2160 GCAGAAGNNNNNNNNNNNNNNNNNNNN------------------------------NNN 2189

Qy          741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
                                                 |||:::|||   ||||||   ||||||
Db         2190 NNNNNNNNNNNNNNNNNNNNNNNNNNNNCTGCTCCAGCTCCCCAGACAGAAGATGCCAAG 2249

Qy          761 AspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAspGlyProProPro 780
                :::
Db         2250 GAG--------------------------------------------------------- 2252

Qy          781 GlyProProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800

Db         2252 ------------------------------------------------------------ 2252

Qy          801 GluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaProProPro 820

Db         2252 ------------------------------------------------------------ 2252

Qy          821 ValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGly 840
                         |||   |||   ||||||                     ::::::   |||
Db         2253 ------CAGAAGTCTGAGGCCGAGGAG---------------------ATCGATGTGGGA 2285

Qy          841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGlu 860
```

```
              :::   ::::::|||   |||:::||||||                           |||
Db       2286 AAGCCAGAGGAGCCCGAGGCCTCTGAGGAG---------------------CCCCCGGAG 2324

Qy        861 ProValLysSerGluCysThrGluGluAlaGluGluGlyProAla---LysGlyLysAsp 879
                 |||||||||:::      ||||||   ||||||   |||      |||   |||
Db       2325 AGTGTAAAGAGTGACCACAAGGAGGAGACCGAGGAAGAGCCTGAAGACAAAGCCAAGGGC 2384

Qy        880 AlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGly 899
                 ||||||   |||      :::|||      ||||||   |||      |||   |||
Db       2385 ACAGAGGCCATTGAAACTGTGTCTGAGGCACCACTTAAGGTGGAG------GAGGCTGGT 2438

Qy        900 SerGlyArgAlaThrThr---AlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
              |||      |||   |||         ||||||||||||   ||||||||||||||||||
Db       2439 AGCAAGGCAGCTGTGACCAAGGGTTCCAGCTCAGGTGCCACCCAGGACAGTGACTCCAGT 2498

Qy        919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
              ||||||||||||||||||||||||||||||   |||||||||||||||   |||||||||
Db       2499 GCCACCTGCAGTGCCGATGAGGTGGACGAACCCGAAGGAGGTGACAAGGGCAGGCTGCTG 2558

Qy        939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
              |||||||||||||||||||||||||||   |||||||||||||||:::   |||||||||
Db       2559 TCACCAAGGCCCAGCCTCCTCACCCCGGCTGGAGATCCCCGGGCCAGTACCTCGCCCCAG 2618

Qy        959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2619 AAGCCGCTGGACCTGAAGCAGCTGAAGCAGCGAGCAGCCGCCATCCCCCCTATCCAGGTC 2678

Qy        979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAla 998
              ||||||||||||||||||||||||||||||      |||   ||||||   ||||||
Db       2679 ACCAAGGTCCATGAGCCCCCCCGGGAGGACACAGTACCCCCAAAGCCAGTTCCCCCTGTG 2738

Qy        999 ProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSer 1018
              |||||||||   |||:::||||||||||||   |||      ||||||   |||   |||
Db       2739 CCTCCACCCACGCAGCACCTACAGCCAGAGGGTGACGTGTCTCAGCAGTCGGGAGGAAGT 2798

Qy       1019 ProArgGlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGlu 1038
              ||||||||||||||||||||||||   |||||||||:::||||||
Db       2799 CCACGTGGCAAGTCCCGCAGCCCAGTGCCTCCTGCCGAGAAAGAGGNNNNNNNNNNNNNN 2858

Qy       1039 AlaGlnLysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValPro 1058

Db       2859 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2918

Qy       1059 ProArgGluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAla 1078

Db       2919 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2978

Qy       1079 ProProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
                       |||||||||||||||||||||||||||:::|||||||||||||||||||||
Db       2979 NNNNNNNGTCACCCGCTGCCTCTGGGCCTCCACGATAGTGCCCGGCCCGTCCTGCCACGT 3038

Qy       1099 ProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeu 1118
              ||||||   ||||||||||||||||||||||||||||||||||||||||||   ||||||
Db       3039 CCCCCC---ATCTCTAACCCCCCACCCCTCATCTCCTCTGCCAAGCATCCCGGCGTACTT 3095

Qy       1119 GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSer 1138
              |||||||||:::|||||||||||||||||||||||||||||||||   ||||||:::|||
Db       3096 GAGAGGCAGCTGGGTGCCATCTCCCAGGGGATGTCAGTCCAGCTTCGTGTGCCTCACTCA 3155

Qy       1139 GluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPro 1158
              ||||||||||||||||||:::||||||:::||||||||||||||||||   :::||||||
Db       3156 GAGCATGCCAAGGCCCCCATGGGCCCTCTCACCATGGGGCTGCCCCTTGCCGTGGACCCT 3215

Qy       1159 LysLysLeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAla 1178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3216 AAGAAGCTGGCCCCCTTCAGTGGAGTGAAGCAGGAACAGTTGTCCCCTCGGGGTCAGGCT 3275

Qy       1179 GlyProProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThr 1198
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       3276 GGGCCACCTGAAAGTCTGGGAGTGCCCACTGCTCAGGAGACCTCTGTGCTGAGAGGGACA 3335

Qy       1199 AlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValPro 1218
              ||||||||||||         ||||||||||||||||||:::||||||||||||
Db       3336 GCACTGGCTCCGCCACCAGTGGAAGCATCACCAAGGGCCTCCCCAGTACCCGGGCTGCA 3395

Qy       1219 SerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspValLeu 1238
                             :::||||||||||||||||||||||||||||||||||||||||||
Db       3396 GACGGCCCC------AGCTACAGAGGCTCTATCACCCACGGCACGCCCGCAGACGTCCTC 3449

Qy       1239 TyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspArgGly 1258
              |||||||||||||||:::||||||:::||||||||||||||||||||||||||||||
Db       3450 TACAAGGGTACCATCAGCAGGATCGTCGGTGAGGACAGCCCAAGTCGCCTTGACCGGGCA 3509

Qy       1259 ArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysLysGlyHisValLeu 1278
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db       3510 CGAGAGGACACCCTGCCCAAGGGCCATGTCATCTATGAGGGCAAGAAAGGCCACGTCCTA 3569

Qy       1279 SerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSer 1298
              |||||||||
Db       3570 TCCTATGAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3629

Qy       1299 GlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVal 1318

Db       3630 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3689

Qy       1319 GlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGlu 1338
                                               ||||||||||||||||||   ||||||
Db       3690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGACTCATGGGCCGCGCCATC---CCTGAG 3746

Qy       1339 ArgHisSerProHisHisLeuLysGluGlnHisHisIleArgGlySerIleThrGlnGly 1358
              :::|||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db       3747 CAGCACAGCCCC---CACCTCAAGGAGCAGCATCACATCCGAGGCTCCATCACGCAAGNN 3803

Qy       1359 IleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLeuLeu 1378

Db       3804 NNNNN-NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3862

Qy       1379 LysArgGluGlyThrProProProProProProSerArgAspLeuThrGluAlaTyrLys 1398

Db       3863 NNNNNN---------------------------NNNNNNNNNNNNNNNNNNNNNNNNNNN 3895

Qy       1399 ThrGlnAlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrVal 1418

Db       3896 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3955

Qy       1419 LysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArgHisThrProGlu 1438

Db       3956 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4015

Qy       1439 LeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLys 1458

Db       4016 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4075

Qy       1459 TyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeuIleGly 1478
                                :::||||||:::|||||||||||||||||||||:::||||||
Db       4076 NNNNNNNNNNNNNCACCCTCCACTGGCACCAAGAAACACGACGTGCGCTCCATCATCGGC 4135

Qy       1479 SerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAla 1498
              ||||||||||||   ||||||   :::||||||||||||:::||||||||||||||||||
Db       4136 AGCCCCGGCCGGCCTTTCCCTGCCCTGCACCCGCTGGACATAATGGCTGATGCCCGGGCA 4195

Qy       1499 LeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSer 1518
              |||||||||||||||||||||||||||||||||||||||   ||||||:::|||   :::
Db       4196 CTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCCGGTCAGGGACCAGCAGTGGTGCA 4255

Qy       1519 GlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGln 1538
              ||||||||||||   |||||||||||||||:::|||||||||||||||||||||||||||
Db       4256 GGGGGCTCCATCACACGTGGGGCTCCAGTCGTCGTGCCTGAACTGGGCAAGCCACGGCAA 4315

Qy       1539 SerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgGlySer 1558
              ||||||||||||||||||||||||||||||||||||      ||||||||||||||||||
Db       4316 AGCCCACTGACTTACGAAGACCACGGGGCACCCTTCACCAGTCACCTGCCACGTGGCTCC 4375

Qy       1559 ProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSerSerLys 1578
              |||||||||   |||||||||||||||||||||||||||
Db       4376 CCTGTGACCACGAGGGAGCCCACGCCACGCCTTCAGGAAGNNNNNNNNNNNNNNNNNNNN 4435
```

```
Qy        1579 AlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSer 1598
Db        4436 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4495

Qy        1599 ThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArgGlyVal 1618
Db        4496 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4555

Qy        1619 SerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSerIlePro 1638
Db        4556 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4615

Qy        1639 ArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsn 1658
                                             |||||||||||||||||||||||||||:::
Db        4616 NNNNNNNNNNNNNNNNNNNNNNCCGCAGCCTACTACCTGCCCCGGCACTTGGCCCCCAGC 4675

Qy        1659 ProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThrAlaAla 1678
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4676 CCCACCTACCCACACCTGTACCCACCTTACCTCATCCGCGGCTACCCTGACACGGCGGCC 4735

Qy        1679 LeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisHisAsn 1698
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4736 CTGGAGAACCGCCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCACCACAAC 4795

Qy        1699 ThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgGluSer 1718
                  |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4796 GCTGCCTCCGCCATGGCCCAGCGTGCTGACATGCTGAGGGGTCTGTCACCGCGAGAGTCC 4855

Qy        1719 SerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValPro 1738
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4856 TCGCTGGCCCTCAATTATGCCGCTGGCCCAAGAGGCATTATCGACCTGTCCCAAGTGCCA 4915

Qy        1739 HisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeu 1758
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        4916 CACCTGCCCGTGCTGGTGCCACCAACGCCAGGCACCCCTGCCACCGCCATCGACCGCCTT 4975

Qy        1759 AlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSer 1778
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        4976 GCCTACCTCCCCACTGCGCCCCCACCCTTCAGCAGCCGCCACAGTAGCTCACCGCTGTCC 5035

Qy        1779 ProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArg 1798
               |||||||||||||||||||||   |||||||||   ||||||||||||||||||||||||
Db        5036 CCAGGAGGCCCCACTCACCTAGCTAAACCAACTGCCACATCTTCATCGGAGCGGGAACGG 5095

Qy        1799 AspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThr 1818
               :::|||:::||||||||||||                  |||||||||||||||||||||
Db        5096 GAACGTGAGCGGGAACGAGAC------------------AAGTCCATCCTCACGTCTACC 5137

Qy        1819 ThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSerGlySerSer 1838
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        5138 ACTACAGTGGAGCATGCACCCATCTGGAGACCTGGTACGGAGCAGAGCAGCGGGGCT--- 5194

Qy        1839 GlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHis 1858
                                    |||   |||||||||||||||||||||:::      |||
Db        5195 ---------------------GGGGGCAGCAGCCGCCCCGCCTCCCACACC------CAC 5227

Qy        1859 GlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeu 1878
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5228 CAGCACTCGCCCATCTCCCCCCGGACCCAGGACGCCTTGCAGCAGAGGCCCAGTGTGCTG 5287

Qy        1879 HisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeu 1898
               |||||||||   |||||||||::::::|||:::|||||||||      ||||||||||||
Db        5288 CACAACACGAGCATGAAGGGCGTGGTCACCTCCGTGGAACCCGGCACGCCCACGGTCCTG 5347

Qy        1899 ArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHis 1918
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5348 AGGTCCACCTCCACCTCTTCGCCTGTCCGCCCAGCTGCCACATTCCCACCTGCCACCCAC 5407

Qy        1919 CysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeu 1938
               |||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        5408 TGCCCACTTGGTGGCACCCTTGAAGGGGTCTACCCTACCCTCATGGAGCCCGTCCTGTTA 5467

Qy        1939 ProLysGluAlaProArgValAlaArgProGluArgProArgAlaAspThrGlyHisAla 1958
               |||||||||      |||||||||||||||||||||||||||   |||   |||||||||
Db        5468 CCCAAGGAGACCTCTCGGGTCGCCCGGCCCGAGCGGCCCCGTGTGGACGCTGGCCATGCC 5527

Qy        1959 PheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGly 1978
               ||||||   |||||||||||||||         ||||||||||||||||||||||||
Db        5528 TTCCTCACCAAACCCCCGGCCCGG---------GAGCCCGCCTCCTCACCCAGCAAGAGC 5578

Qy        1979 SerGluProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrPro 1998
               ||||||||||||   |||   ||||||   |||   |||      |||||||||||||||
Db        5579 TCCGAGCCCCGATCCCTAGCACCCCCCAGCTCCAGCCACACAGCCATCGCCCGCACCCCA 5638

Qy        1999 AlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAla 2018
               ||||||:::||||||||||||||||||||||||||||||||||||   |||   ||||||
Db        5639 GCAAAGAGCCTTGCACCCCACCATGCCAGTCCGGACCCGCCGGCG---CCCACCTCGGCC 5695

Qy        2019 SerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeu 2038
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5696 TCAGATCTGCACCGAGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAATTGGAACTC 5755

Qy        2039 ArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerPro 2058
               |||||||||
Db        5756 CGTTCTCTGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5815

Qy        2059 ValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAsp 2078
Db        5816 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5875

Qy        2079 LysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGly 2098
                                                         |||:::||||||
Db        5876 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGCCCCATGAAGCTCAGCGCG 5935

Qy        2099 GluAlaAlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerPro 2118
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5936 GAGGCTGCCCATCTCCCACATCTGCGGCCACTGCCCGAGAGCCAGCCCTCATCCAGCCCA 5995

Qy        2119 LeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHis 2138
               |||||||||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        5996 CTCCTCCAGACTGCCCCAGGCATCAAAGGTCACCAGAGGGTGGTCACCCTGGCTCAGCAC 6055

Qy        2139 IleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaPro 2158
               ||||||
Db        6056 ATCAGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6115

Qy        2159 LeuProAlaProLeuTyrSerPheProGlyAlaSerCysProValLeuAspLeuArgArg 2178
Db        6116 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6175

Qy        2179 ProProSerAspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerPro 2198
Db        6176 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6235

Qy        2199 HisSerGluGlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGly 2218
                                    ||||||||||||:::||||||||||||||||||
Db        6236 NNNNNNNNNNNNNNNNNNNNNGTCCCAGAACCCAGCAAAACATCGGTCCTGGGCAGCAGT 6295

Qy        2219 GluAspGlyIleGluProValSerProProGluGlyMetThrGluProGlyHisSerArg 2238
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        6296 GAGGATGCCATTGAGCCTGTGTCCCCACCAGAGGGCATGACTGAGCCAGGACATGCTCGG 6355

Qy        2239 SerAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGly 2258
               |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        6356 AGCGCTGTGTACCCACTGCTGTATCGAGACGGGGAACAGGGCGAGCCCAGCAGGATGGGC 6415

Qy        2259 SerLysSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSer 2278
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6416 TCTAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGTAAGCTGACTGAGAGC 6475

Qy        2279 AsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsn 2298
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        6476 AACTCCGCCATGGTGAAGTCGAAGAAGCAGGAGATCAACAAGAAACTCAACACCCACAAC 6535

Qy        2299 ArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAla 2318
```

```
            ||||||||||||||||||
Db     6536 CGGAACGAGCCAGAATACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6595

Qy     2319 IleThrGlyThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThr 2338

Db     6596 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6655

Qy     2339 AsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGlu 2358

Db     6656 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6715

Qy     2359 GluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuPro 2378

Db     6716 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6775

Qy     2379 AlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGly 2398

Db     6776 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGT 6835

Qy     2399 GlyGlyGlyLysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAla 2418
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6836 GGAGGTGGGAAAGCCAAGGTCTCTGGCAGACCTAGCAGCCGAAAAGCCAAGTCGCCAGCA 6895

Qy     2419 ProGlyLeuAlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAsp 2438
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6896 CCAGGCCTAGCGTCCGGAGACCGACCCCCTTCTGTCTCCTCAGTACACTCAGAGGGGGAC 6955

Qy     2439 CysAsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAla 2457
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     6956 TGCAATCGCCGAACACCACTCACCAACCGTGTGTGGGAGGACCGGCCCTCATCTGCA 7012

RESULT 3
AY412687
LOCUS       AY412687                6836 bp    DNA     linear   GSS 16-DEC-2003
DEFINITION  Pan troglodytes NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY412687
VERSION     AY412687.1  GI:39768652
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 6836)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2  (bases 1 to 6836)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES             Location/Qualifiers
     source          1. .6836
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene            <1. .>6836
                     /gene="NCOR2"
                     /locus_tag="HCM4640"
ORIGIN

Alignment Scores:
Pred. No.:                0             Length:          6836
Score:                    6533.00       Matches:         1407
Percent Similarity:       61.81%        Conservative:    1
Best Local Similarity:    61.76%        Mismatches:      870
Query Match:              49.44%        Indels:          0
DB:                       9             Gaps:            0

US-09-522-753-5 (1-2517) x AY412687 (1-6836)

Qy       27 ProValGlnIleAlaArgThrHisThrAspValGlyLeuLeuGluTyrGlnHisHisSer 46
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2 CCAGTGCAGATTGCCCGGACGCACACGGACGTCGGGCTCCTGGAGTACCAGCACCACTCC 61

Qy       47 ArgAspTyrAlaSerHisLeuSerProGlySerIleIleGlnProGlnArgArgArgPro 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       62 CGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAGCCCCAGCGGCGGAGGCCC 121

Qy       67 SerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgPro 86
            |||||||||||||||||||||||||||||||||||
Db      122 TCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGNNNNNNNNNNNNNNNNNNNNNNNNN 181

Qy       87 GluSerHisSerTyrLeuProGluLeuGlyLysSerGluMetGluPheIleGluSerLys 106

Db      182 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 241

Qy      107 ArgProArgLeuGluLeuLeuProAspProLeuLeuArgProSerProLeuLeuAlaThr 126

Db      242 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 301

Qy      127 GlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThrGlyLysLeuGlu 146
                                                   |||||||||||||||||||||
Db      302 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGCCTGACGGGCAAGCTGGAA 361

Qy      147 ProValSerProProSerProProHisThrAspProGluLeuGluLeuValProProArg 166
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      362 CCGGTGTCTCCCCCCAGCCCCCCGCACACCGACCCTGAGCTGGAGCTGGTGCCGCCGCGG 421

Qy      167 LeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGluIleThrMetVal 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      422 CTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGACCGAGAGATCACCATGGTA 481

Qy      187 GluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGluGluGluAlaAlaLys 206
            |||||||||||||||||||||||||||||||||         ||||||||||||   |||
Db      482 GAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGNNNNNNNTGGAGGAGGAGGCTNCCAAG 541

Qy      207 ProProGluProGluLysProValSerProProProIleGluSerLysHisArgSerLeu 226
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      542 CCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAGTCGAAGCACCGCAGCCTG 601

Qy      227 ValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAlaHisArgIleLeuGlu 246
            |||||||||||||||||||||||||||
Db      602 GTGCAGATCATCTACGACGAGAACCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 661

Qy      247 GlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyr 266
                                    ||||||||||||||||||||||||||||||||||||
Db      662 NNNNNNNNNNNNNNNNNNNNNNNNCCGCTGTACAACCAGCCCTCCGACACCCGGCAGTAT 721

Qy      267 HisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeuIleLeuTyrPheLysArg 286
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      722 CATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTAATCTTGTACTTCAAGAGG 781

Qy      287 ArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAspGlnLeuMet 306
            ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db      782 AGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAGCGCTATGACCAGCTCATG 841

Qy      307 GluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgArgAlaLysGlu 326
            ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      842 GAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCCCGGCGGCGGGCCAAGGAG 901

Qy      327 SerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLysGlnArgGluLeu 346
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      902 AGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATCCGCAAGCAGCGCGAGCTG 961
```

```
Qy         347 GlnGluArgMetGlnSerArgValGlyGlnArgGlySerGlyLeuSerMetSerAlaAla 366
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         962 CAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGGTTGTCCATGTCGGCTGCC 1021

Qy         367 ArgSerGluHisGluValSerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGlu 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1022 CGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCAGAGCAGGAGAACCTGGAG 1081

Qy         387 LysGlnMetArgGlnLeuAlaValIleProProMetLeuTyrAspAlaAspGlnGlnArg 406
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1082 AAGCAGATGCGCCAGCTGGCCGTGATCCCGCCTATGCTGTACGACGCCGACCAGCAGCGC 1141

Qy         407 IleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMetLysValTyrLysAspArg 426
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1142 ATCAAGTTCATCAACATGAACGGGCTTATGGCTGACCCCATGAAGGTGTACAAAGACCGC 1201

Qy         427 GlnValMetAsnMetTrpSerGluGlnGluLysGluThrPheArgGluLysPheMetGln 446
               ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1202 CAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTCCGGGAGAANNNNNNNNNN 1261

Qy         447 HisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArgLysThrValAlaGluCys 466
                                                               ||||||||||||
Db        1262 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTGGCTGAGTGC 1321

Qy         467 ValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLysSerLeuValArgArgSer 486
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1322 GTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAGAGCCTGGTGAGACGGAGC 1381

Qy         487 TyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 506
               ||||||||||||||||||||||||
Db        1382 TATCGGCGCCGCGGCAAGAGCCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1441

Qy         507 GlnGlnGlnGlnProMetProArgSerSerGlnGluGluLysAspGluLysGluLysGlu 526

Db        1442 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1501

Qy         527 LysGluAlaGluLysGluGluGluLysProGluValGluAsnAspLysGluAspLeuLeu 546

Db        1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561

Qy         547 LysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValAlaSer 566
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1562 NNGGAGAAGACAGACGACACCTCAGGGGAGGACAACGACGAGAAGGAGGCCGTGGCCTCC 1621

Qy         567 LysGlyArgLysThrAlaAsnSerGlnGlyArgArgLysGlyArgIleThrArgSerMet 586
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1622 AAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAAGGCCGCATCACCCGCTCAATG 1681

Qy         587 AlaAsnGluAlaAsnSerGluGluAlaIleThrProGlnGlnSerAlaGluLeuAlaSer 606
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1682 GCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAGCAGAGCGCCGAGCTGGCCTCC 1741

Qy         607 MetGluLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGly 626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1742 ATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAAATGGAAACAGCCAAGAAAGGT 1801

Qy         627 LeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrVal 646
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1802 CTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGGATGGTGGGCTCCAAGACTGTG 1861

Qy         647 SerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeuAspGluIleLeu 666
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1862 TCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGGCAGAACCTCGATGAGATCTTG 1921

Qy         667 GlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAla 686
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1922 CAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCACGGAGGAAGAAGAAGAAAGCG 1981

Qy         687 ProAlaAlaAlaSerGluGluAlaAlaPheProProValValGluAspGluGluMetGlu 706
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1982 CCGGCGGCGGCCAGCGAGGAGGCTGCATTCCCGCCCGTGGTGGAGGATGAGGAGATGGAG 2041

Qy         707 AlaSerGlyValSerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAla 726
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db        2042 GCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGAGGAGGCTGAAGNNNNNNNNNNN 2101

Qy         727 SerGlyAsnGluValProArgGlyGluCysSerGlyProAlaThrValAsnAsnSerSer 746

Db        2102 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2161

Qy         747 AspThrGluSerIleProSerProHisThrGluAlaAlaLysAspThrGlyGlnAsnGly 766

Db        2162 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2221

Qy         767 ProLysProProAlaThrLeuGlyAlaAspGlyProProProGlyProProThrProPro 786

Db        2222 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2281

Qy         787 ArgArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaPro 806

Db        2282 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2341

Qy         807 ThrProProProAlaProProSerProSerAlaProProProValValProLysGluGlu 826

Db        2342 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2401

Qy         827 LysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLysProPro 846
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2402 NAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAGGAGGGGGAGGAGCAGAAGCCCCCC 2461

Qy         847 AlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCys 866
               |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2462 GCGNNNGAGGAGCTGGCAGTGGACACAGGGAAGGCCGAGGAGCCCGTCAAGAGCGAGTGC 2521

Qy         867 ThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThr 886
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2522 ACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAGGACGCCGAGGCCGCTGAGGCTACG 2581

Qy         887 AlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAla 906
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2582 GCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGCGGGAGCGGCAGGGCCACCACAGCC 2641

Qy         907 LysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluVal 926
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2642 AAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGTGCCACCTGCAGTGCAGACGAGGTG 2701

Qy         927 AspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThr 946
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2702 GATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCAAGGCCCAGCCTCCTCACC 2761

Qy         947 ProThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeu 966
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2762 CCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCACTGGACCTGAAGCAGCTG 2821

Qy         967 LysGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArg 986
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2822 AAGCAGCGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAAGTCCATGAGCCCCCCCGG 2881

Qy         987 GluAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGln 1006
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2882 GAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCGCCACCGCAAAACCTGCAG 2941

Qy        1007 ProGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerPro 1026
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2942 CCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGGGGCAAGAGCAGGAGCCCG 3001

Qy        1027 AlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspPro 1046
               |||||||||||||||||||||
Db        3002 GCACCCCCCGCCGACAAGGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3061

Qy        1047 ProCysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSer 1066

Db        3062 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3121

Qy        1067 ProHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeu 1086
                                                            |||||||||||||||
Db   3122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCACCCACTGCCCCTG 3181

Qy   1087 GlyLeuHisAspThrAlaArgProValLeuProArgProProThrIleSerAsnProPro 1106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3182 GGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCT 3241

Qy   1107 ProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSer 1126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3242 CCCCTTATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCC 3301

Qy   1127 GlnGlyMetSerValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGly 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3302 CAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGC 3361

Qy   1147 ProValThrMetGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGly 1166
          |||||||||||||||||||||||||||||||||||||||||||||||||||      |||
Db   3362 CCCGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTNNGCGGA 3421

Qy   1167 ValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyVal 1186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3422 GTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGAGTG 3481

Qy   1187 ProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGly 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3482 CCCACGGCCCAGGAGGCATCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGA 3541

Qy   1207 SerIleThrLysGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArg 1226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3542 AGCATCACCAAAGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGC 3601

Qy   1227 GlySerIleThrHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIle 1246
          |||||||||||||||   ||||||||||||||||||      ||||||||||||||||||
Db   3602 GGCTCCATCACCCACNNNACGCCAGCTGACGTCCTGTANNNNGGCACCATCACCAGGATC 3661

Qy   1247 IleGlyGluAspSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGly 1266
          |||||||||||||||      ||||||      |||||||||||||||||||||||||||
Db   3662 ATCGGCGAGGACAGCCNNNGTCGCTTGGANNGCGGCCGGGAGGACAGCCTGCCCAAGGGC 3721

Qy   1267 HisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerVal 1286
          |||||||||||||||||||||||||||||||||||||||||||||
Db   3722 CACGTCATCTACGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGNNNNNNNNNNNNNN 3781

Qy   1287 ThrGlnCysSerLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAla 1306

Db   3782 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3841

Qy   1307 ProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSer 1326

Db   3842 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3901

Qy   1327 IleGluGlyLeuMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLys 1346
                   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3902 NNNNNNNGTCTCATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAA 3961

Qy   1347 GluGlnHisHisIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAla 1366
          |||||||||||||||||||||||||||||||||
Db   3962 GAGCAGCACCACATCCGCGGGTCCATCACACAAGNNNNNNNNNNNNNNNNNNNNNNNNNN 4021

Qy   1367 GlnGluAspTyrLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProPro 1386

Db   4022 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4081

Qy   1387 ProProProSerArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLys 1406

Db   4082 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4141

Qy   1407 LeuLysProAlaHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHis 1426

Db   4142 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4201

Qy   1427 GluIleProArgGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeu 1446

Db   4202 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4261

Qy   1447 LysGluGlySerIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThr 1466

Db   4262 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4321

Qy   1467 GlySerLysLysHisAspValArgSerLeuIleGlySerProGlyArgThrPheProPro 1486
                            ||||||||||||||||||||||||||||||||||||||||||
Db   4322 NNNNNNNNNNNNNNNNACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCC 4381

Qy   1487 ValHisProLeuAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGlu 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4382 GTGCACCCGCTGGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAG 4441

Qy   1507 SerLeuLysSerArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAla 1526
          |||||||||||||||||||||         ||||||||||||||||||||||||||||||
Db   4442 AGCCTGAAGAGCCGGCCAGGCNNNNNNNNNAGCTCGGGGGGCTCCATTGCGCGCGGCGCC 4501

Qy   1527 ProValIleValProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHis 1546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4502 CCGGTCATTGTGCCTGAGCTGGGCAAGCCGCGGCAGAGCCCCCTGACCTATGAGGACCAC 4561

Qy   1547 GlyAlaProPheAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThr 1566
          |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db   4562 GGGGCACCCTTTGCCGGCCACCTCCCGCGAGGTTCGCCCGTGACCACGCGGGAGCCCACG 4621

Qy   1567 ProArgLeuGlnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThr 1586
          |||||||||||||||
Db   4622 CCGCGCCTGCAGGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4681

Qy   1587 SerThrProArgGluIleAlaLysSerProHisSerThrValProGluHisHisProHis 1606

Db   4682 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4741

Qy   1607 ProIleSerProTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSer 1626

Db   4742 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4801

Qy   1627 HisIleProLeuAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAla 1646

Db   4802 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCC 4861

Qy   1647 AlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrPro 1666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4862 GCTGCCTACTACCTGCCCCGGCACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCA 4921

Qy   1667 ProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIle 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4922 CCCTACCTCATCCGCGGCTACCCCGACACGGCAGCGCTGGAGAACCGGCAGACCATCATC 4981

Qy   1687 AsnAspTyrIleThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArg 1706
          ||||||||||||||||||||||||||||||||||||   ||||||   ||||||||||||
Db   4982 AATGACTACATCACCTCGCAGCAGATGCACCACAACGCGGCCACCACCATGGCCCAGCGA 5041

Qy   1707 AlaAspMetLeuArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAla 1726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5042 GCTGACATGCTGAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCG 5101

Qy   1727 GlyProArgGlyIleIleAspLeuSerGlnValProHisLeuProValLeuValProPro 1746
          |||||||||
Db   5102 GGTCCCCGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5161

Qy   1747 ThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGln 1766

Db   5162 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5221

Qy   1767 ProPheSerSerArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThr 1786

Db   5222 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5281

Qy   1787 LysProThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArg 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5282 AAACCAACCACCACATCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGNNNNNNN 5341
```

```
Qy       1807 AspArgGluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIle 1826
                    ||||||||||||   |||||||||||||||      ||||||||||||||||||
Db       5342 NNNNNNGAGCGGGAAAAGNCCATCCTCACGTCCACNNNGNNNGTGGAGCACGCCCCCATC 5401

Qy       1827 TrpArgProGlyThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGly 1846
              |||         ||||||||||||||||||||||||||||||||||||||||||||||||
Db       5402 TGGAGNNNNNGTACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGTGGGGGTGGGGGC 5461

Qy       1847 SerSerSerArgProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArg 1866
              |||||||||||||||||||||||||||      ||||||||||||||||||||||||
Db       5462 AGCAGCAGCCGCCCCGCCTCCCACTCCCANNCCCACCAGCACTCGCCCATCTCCCCTNNN 5521

Qy       1867 ThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIle 1886

Db       5522 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5581

Qy       1887 IleThrAlaValGluProSerLysProThrValLeuArgSerThrSerThrSerSerPro 1906

Db       5582 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5641

Qy       1907 ValArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAsp 1926

Db       5642 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5701

Qy       1927 GlyValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAla 1946

Db       5702 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5761

Qy       1947 ArgProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArg 1966

Db       5762 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5821

Qy       1967 SerGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValPro 1986
                                                           |||||||||||||||
Db       5822 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGGCCCCTNGTGCCC 5881

Qy       1987 ProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHis 2006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5882 CCCGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCAC 5941

Qy       2007 AlaSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThr 2026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5942 GCCAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACT 6001

Qy       2027 GlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySer 2046
              |||||||||||||||||||||||||||||||||||||||||||||
Db       6002 CAAAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGNNNNNNNNNNNNNN 6061

Qy       2047 SerTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHis 2066

Db       6062 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6121

Qy       2067 AspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeu 2086

Db       6122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6181

Qy       2087 ArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeu 2106
                                ||||||||||||||||||||||||||||||||||||||||||
Db       6182 NNNNNNNNNNNNNNNNGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTG 6241

Qy       2107 ArgProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyVal 2126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6242 CGGCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTC 6301

Qy       2127 LysGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAsp 2146
              ||||||||||||||||||||||||||||||
Db       6302 AAAGGTCACCAGCGGGTGGTCACCCTGGCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6361

Qy       2147 TyrThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhe 2166

Db       6362 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6421

Qy       2167 ProGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuPro 2186

Db       6422 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 6481

Qy       2187 ProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSer 2206
                                                                       |||
Db       6482 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCT 6541

Qy       2207 ProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSer 2226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6542 CCAGAGCCAAACAAGACGTCAGTCTTGGGTGGCGGTGAGGACGGTATTGAACCTGTGTCC 6601

Qy       2227 ProProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyr 2246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6602 CCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTAC 6661

Qy       2247 ArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSer 2266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6662 CGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGC 6721

Qy       2267 GlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLys 2286
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6722 CAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAG 6781

Qy       2287 LysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyr 2304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6782 AAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATAC 6835

RESULT 4
BC033087
LOCUS       BC033087                2000 bp    mRNA    linear   HTC 04-MAR-2003
DEFINITION  Homo sapiens, Similar to nuclear receptor co-repressor 2, clone
            IMAGE:5016291, mRNA.
ACCESSION   BC033087
VERSION     BC033087.1  GI:21619972
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2000)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 43 Row: i Column: 21
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 5454073
            This clone has the following problem: retained intron.
FEATURES             Location/Qualifiers
     source          1. .2000
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5016291"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
ORIGIN

Alignment Scores:
Pred. No.:              4.36e-140      Length:          2000
Score:                  2566.00        Matches:         503
Percent Similarity:     90.72%         Conservative:    15
Best Local Similarity:  88.09%         Mismatches:      31
Query Match:            19.42%         Indels:          22
DB:                     3              Gaps:            1

US-09-522-753-5 (1-2517) x BC033087 (1-2000)

Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        273 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 332

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 392

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 452

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 512

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 572

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 632

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 692

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 752

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 812

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 872

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 932

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        933 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 992

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        993 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 1052

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1053 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 1112

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       1113 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1172

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db       1173 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1232

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1233 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1292

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1293 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1352

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1412

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1413 GAGCAGGAGAACCTGGAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1472

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1473 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1532

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1533 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1592

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1593 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1652

Qy        461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1653 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1712

Qy        481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
              |||||||||||||||||||||||||||||||||||||||            :::   |||
Db       1713 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCTGCTGCTGCTAAAGTTCCAG 1772

Qy        501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGluLys 520
              |||::::::         :::   ::::::      ||||||               |||
Db       1773 CAAAAAAGATCACCGCCGCGAGTAAAAAGGCTCCAGCCCAGAAGGTTCCTGCCCAGAAAG 1832

Qy        521 AspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGluAsn 540
                 :::      |||::::::      :::|||               :::::::::
Db       1833 CCACAGGCCAGAAAGCAGCGCCTGCTCCAAAAGCTCAGAAGGGTCAAAAAGCTCCAGCCC 1892

Qy        541 AspLysGlu--------------------------------------------------- 543
                 |||
Db       1893 AGAAAGCACCTGCTCCAAAGGCATCTGGCAAGAAAGCATAAGTGGCAATCATAAAAAGTA 1952

Qy        544 ---------------AspLeuLeuLysGluLys 549
                             ||||||||||||:::|||
Db       1953 ATAAAGGTTCTTTTTGACCTGTTAAAAAAAAAA 1985

RESULT 5
BC020427
LOCUS       BC020427                1362 bp    mRNA    linear   HTC 03-JAN-2002
DEFINITION  Homo sapiens, clone IMAGE:4179307, mRNA.
ACCESSION   BC020427
VERSION     BC020427.1  GI:18043012
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 1362)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm .tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M.,Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gunaratne, P., Yoon, V., Kowis, C., Martin, R.,
            Lawrence, S., Richards, S., Gibbs,R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 12 Row: i Column: 23
            This clone has the following problem: no polyA-tail.
FEATURES             Location/Qualifiers
     source          1. .1362
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4179307"
                     /tissue_type="Brain, anaplastic oligodendroglioma with
                     1p/19q loss"
                     /clone_lib="NCI_CGAP_Brn67"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-122     Length:          1362
Score:                  2264.00       Matches:         431
Percent Similarity:     99.54%        Conservative:    0
Best Local Similarity:  99.54%        Mismatches:      2
Query Match:            17.13%        Indels:          0
DB:                     3             Gaps:            0

US-09-522-753-5 (1-2517) x BC020427 (1-1362)

Qy       1676 ThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMet 1695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 ACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATG 122

Qy       1696 HisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerPro 1715
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 CACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCC 182

Qy       1716 ArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSer 1735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGGCATCATCGACCTGTCC 242

Qy       1736 GlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMet 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCACCCCAGCCACCGCCATG 302

Qy       1756 AspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSer 1775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAGCCGCCACAGCAGCTCC 362

Qy       1776 ProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGlu 1795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCACCACGTCCTCGTCCGAG 422

Qy       1796 ArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgGluLysSerIleLeu 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 CGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAGTCCATCCTC 482

Qy       1816 ThrSerThrThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSerSer 1835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 ACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGGTACAGAGCAGAGCAGC 542

Qy       1836 GlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSer 1855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 GGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCGCCCCGCCTCCCACTCC 602

Qy       1856 HisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgPro 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGCCCTCCAGCAGAGACCC 662

Qy       1876 SerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysPro 1895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        663 AGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGTGGAGCCCAGCACGCCC 722

Qy       1896 ThrValLeuArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProPro 1915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ACGGTCCTGAGGTCCACCTCCACCTCCTCACCCGTTCGCCCGGCTGCCACATTCCCACCT 782

Qy       1916 AlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluPro 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GCCACCCACTGCCCACTGGGCGGCACCCTCGATGGGGTCTACCCTACCCTCATGGAGCCC 842

Qy       1936 ValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAlaAspThr 1955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 GTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCGGCCAGAGCGGCCCCGAGCAGACACC 902

Qy       1956 GlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSerSerPro 1975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 GGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTCCGGGCTGGAGCCCGCCTCCTCCCCC 962

Qy       1976 SerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThrIleAla 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 AGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCCTGTCTCTGGCCACGCCACCATCGCC 1022

Qy       1996 ArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProProAlaProPro 2015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1023 CGCACCCCTGCGAAGAACCTCGCACCTCACCACGCCAGCCCGGACCCGCCGGCGCCACCT 1082

Qy       2016 AlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSerIleGlnGlu 2035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCAAAGTAAACCCTTTTCCATCCAGGAA 1142

Qy       2036 LeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluPro 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 CTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAGCTACAGCCCCGAAGGGGTGGAGCCC 1202

Qy       2056 ValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGlu 2075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1203 GTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGACAAGGGGCTCCCCAAGCACCTGGAA 1262

Qy       2076 GluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGlyProValLys 2095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1263 GAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCGGCCCAAGCAGCCAGGCCCCGTGAAG 1322

Qy       2096 LeuGlyGlyGluAlaAlaHisLeuProHisLeuArgPro 2108
              |||||||||||||||||||||||||||||||||||||||
Db       1323 CTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCGGCCG 1361

RESULT 6
BC026028
LOCUS       BC026028                1966 bp    mRNA    linear   HTC 22-MAR-2002
DEFINITION  Homo sapiens, Similar to nuclear receptor co-repressor 1, clone
            IMAGE:3836580, mRNA.
ACCESSION   BC026028
VERSION     BC026028.1  GI:19684000
KEYWORDS    HTC.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1966)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: b Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 5454137
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1966
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3836580"
                     /tissue_type="Ovary, adenocarcinoma"
                     /clone_lib="NIH_MGC_9"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
ORIGIN

Alignment Scores:
Pred. No.:              5.95e-86        Length:         1966
Score:                  1643.50         Matches:        343
Percent Similarity:     73.35%          Conservative:   56
Best Local Similarity:  63.05%          Mismatches:     98
Query Match:            12.44%          Indels:         47
DB:                     3               Gaps:           10

US-09-522-753-5 (1-2517) x BC026028 (1-1966)

Qy          16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
               :::   ||||||||||||||||||:::   |||               |||
Db         392 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 451

Qy          36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
               :::      :::   :::|||:::      |||               ||||||   |||
Db         452 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 511

Qy          54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                 |||      ||||||||||||||||||||||||||||||
Db         512 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 571

Qy          74 ProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSerHisSerTyrLeuPro 93
               ||||||::::::|||   ||||||         |||               :::   |||
Db         572 CCAGGTTCTGACAGGCCTCAAGAA------AGGAGAACTAGTTATGAACCGTTTCATCCA 625

Qy          94 GluLeuGlyLysSerGluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
                              :::   :::   :::||||||||||||||||||||||||   :::
Db         626 GGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCGACCACGTCTGGAACAGGTT 685

Qy         114 ProAsp---------------------------------ProLeuLeuArgProSerPro 122
                  |||                                 ||||||:::   |||   |||
Db         686 TCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCCTTTAGTGCACCCGCTGCCA 745

Qy         123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
                           |||      |||   |||   |||      ||||||   :::
Db         746 ---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAAGAAGGATCCAGCATTCGGA 793

Qy         143 GlyLysLeuGlu---ProValSerProProSerProProHisThrAspProGluLeuGlu 161
               ||||||   |||   |||   ||||||   |||                  :::
Db         794 GGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACCATGTGGAGATGATCAAAAT 853

Qy         162 LeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArg 181
                     |||   :::||||||||||||||||||||||||:::||||||||||||||||||
Db         854 GCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAGTATGGATCGTGTAGATCGA 913

Qy         182 GluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGlu 201
               ||||||   ||||||||||||||||||   ||||||||||||||||||||||||||||||
Db         914 GAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAAGAAACAACAACAGCTTGAA 973

Qy         202 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGluSer 221
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db         974 GAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTCCCCTCCTCCTGTGGAGCAG 1033

Qy         222 LysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAlaAla 241
               ||||||||||||:::|||||||||||||||||||||||||||||||||||||||   |||
Db        1034 AAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCGGAAAAAAGCAGAAGAAGCT 1093

Qy         242 HisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnProSer 261
               |||:::|||   |||||||||||||||:::||||||||||||||||||||||||||||||
Db        1094 CATAAAATTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCCACTGTATAACCAGCCATCA 1153

Qy         262 AspThrArgGlnTyrHisGluAsnIleLysIle--------------------------- 272
               ||||||:::   ||||||||||||||||||
Db        1154 GATACCAAGGTGTACCATGAGAACATCAAGACTGGAGTACCTGCAAGGCGCATGATGAAA 1213

Qy         273 AsnGlnAlaMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLys 292
               ||||||   |||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        1214 AACCAGGTGATGAGGAAAAAACTCATTTATTTTTAAAAGAAGAAATCATGCAAGAAAA 1273

Qy         293 GlnTrpLysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLys 312
               |||   :::||||||   ||||||||||||||||||||||||||||||   |||||||||
Db        1274 CAAAGGGAACAAAAAATCTGCCAGCGTTATGATCAGCTCATGGAGGCATGGGAGAAAAAA 1333

Qy         313 ValGluArgIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyr 332
               |||:::||||||||||||||||||||||||:::|||||||||||||||   |||||||||
Db        1334 GTGGACAGAATAGAAAATAATCCTCGGAGGAAAGCTAAAGAAAGCAAAACAAGGGAATAC 1393

Qy         333 TyrGluLysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSer 352
               |||||||||||||||||||||||||||||||||||||||   |||||||||   |||
Db        1394 TATGAAAAGCAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG--- 1450

Qy         353 ArgValGlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluVal 372
               ||||||||||||||||||:::|||||||||   :::   ||||||||||||||||||:::
Db        1451 CGAGTTGGGCAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATT 1510

Qy         373 SerGluIleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeu 392
               ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        1511 TCTGAAATTATTGATGGGCTCTCTGAGCAGGAGAATAATGAGAAACAAATGCGGCAGCTC 1570

Qy         393 AlaValIleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMet 412
               :::|||||||||||||||::::::||||||:::|||:::|||:::|||||||||||||||
Db        1571 TCTGTGATTCCACCTATGATGTTTGATGCAGAACAAAGACGAGTCAAGTTCATTAACATG 1630

Qy         413 AsnGlyLeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrp 432
               ||||||||||||   ||||||||||||||||||||||||||||||   ||||||:::|||
Db        1631 AATGGGCTTATGGAGGACCCTATGAAAGTGTATAAAGATAGGCAGTTTATGAATGTTTGG 1690

Qy         433 SerGluGlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGly 452
```

```
              ::::::   |||||||||   |||::::::||||||:::|||||||||||||||||||||
Db       1691 ACTGACCATGAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGA 1750

Qy        453 LeuIleAlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeu 472
              ||||||||||||:::||||||||||||:::|||   :::|||||||||||||||||||||
Db       1751 CTAATTGCATCATACTTGGAGAGGAAGAGTGTTCCTGATTGTGTTTTGTATTACTATTTA 1810

Qy        473 ThrLysLysAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGly 491
              ||||||||||||||||||||||||:::||||||||||||:::|||   :::|||||||||
Db       1811 ACCAAGAAAAATGAGAATTATAAAGCCCTCGTCAGAAGGAATTATGGGAAACGCAGAGGC 1870

Qy        492 LysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPro 511
              ::::::||||||
Db       1871 AGAAACCAGCAA------------------------------------------------ 1882

Qy        512 MetProArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLys 531
              :::   |||   |||||||||||||||   |||:::::::::|||   :::|||||||||
Db       1883 ATTGCTCGACCCTCGCAAGAAGAAAAAGTAGAAGAAAAAGAAGAGGATAAAGCAGAAAAA 1942

Qy        532 GluGluGluLys 535
                 |||:::|||
Db       1943 ACAGAAAAAAAA 1954

RESULT 7
BC035748
LOCUS       BC035748                1850 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens nuclear receptor co-repressor 1, mRNA (cDNA clone
            IMAGE:5749816), with apparent retained intron.
ACCESSION   BC035748
VERSION     BC035748.1  GI:23958885
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1850)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1850)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 79 Row: m Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 22538460
            This clone has the following problem: retained intron.
FEATURES             Location/Qualifiers
     source          1. .1850
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5749816"
                     /tissue_type="Pancreas, Spleen, adult pooled"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:              1.97e-85        Length:         1850
Score:                  1634.00         Matches:        346
Percent Similarity:     74.35%          Conservative:   57
Best Local Similarity:  63.84%          Mismatches:     96
Query Match:            12.36%          Indels:         43
DB:                     3               Gaps:           12

US-09-522-753-5 (1-2517) x BC035748 (1-1850)

Qy         16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
              :::   ||||||||||||||||||:::   |||               |||
Db        272 CAAAGTCGTTATCCTCCTCACTCTGTCCAGTATACATTTCCCAACACCCGCCACCAGCAG 331

Qy         36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
              :::      :::   :::|||:::      |||               ||||||   |||
Db        332 GAGTTCGCAGTCCCTGATTATCGTTCCTCTCATCTTGAAGTGAGTCAGGCATCACAGCTT 391

Qy         54 SerProGlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGln 73
                                |||      ||||||||||||||||||||||||||||||
Db        392 TTGCAGCAACAGCAGCAGCAACAGCTTCGAAGGCGACCTTCCTTGCTTTCAGAATTTCAC 451

Qy         74 ProGlyAsnGluArgSerGlnGluLeu---His-LeuArgProGluSerHis------Se 90
              ||||||::::::            |||       :::||||||:::            ||
Db        452 CCAGGTTCTGACAGTGTTTATTTCTTTGTATGTATTAGGCCTCAAGAAAGGAGAACTAG 511

Qy         90 rTyrLeuPro---GluLeuGlyLysSerGluMetGluPhe------IleGluSerLysAr 107
              ||||   |||         |||   |||   ::::::         :::|||||||||||
Db        512 TTATGAACCGTTTCATCCAGGCCCATCCCCAGTGGATCATGATTCACTGGAATCGAAGCG 571

Qy        107 gProArgLeuGluLeuLeuProAsp---------------------------------Pr 116
              |||||||||||||   :::   |||                                 ||
Db        572 ACCACGTCTGGAACAGGTTTCTGATTCTCATTTTCAGCGTGTCAGTGCTGCGGTTTTGCC 631

Qy        116 oLeuLeuArgProSerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuTh 136
              ||||:::   |||   |||            |||      |||   |||   |||
Db        632 TTTAGTGCACCCGCTGCCA---------GAAGGGCTGAGGGCT---TCTGCAGATGCTAA 679

Qy        136 rLysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProProHi 155
               ||||||   :::      ||||||   |||   |||   ||||||   |||
Db        680 GAAGGATCCAGCATTCGGAGGCAAACATGAAGCTCCATCCTCTCCAATTTCGGGGCAACC 739
```

```
Qy          155 sThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAs 175
                          :::            |||   :::||||||||||||||||||||||||::
Db          740 ATGTGGAGATGATCAAAATGCTTCACCTTCAAAACTCTCAAAGGAAGAGTTAATACAGAG 799

Qy          175 nMetAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLy 195
                :||||||||||||||||||||||||      |||||||||||||||   |||||||||||
Db          800 TATGGATCGTGTAGATCGAGAAATTGCAAAAGTAGAACAGCAGATCCTTAAACTGAAAAA 859

Qy          195 sLysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSe 215
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          860 GAAACAACAACAGCTTGAAGAAGAGGCAGCTAAACCTCCTGAGCCTGAGAAGCCCGTGTC 919

Qy          215 rProProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnAr 235
                ||||||||||:::|||   ||||||||||||:::||||||||||||||||||||||||||
Db          920 CCCTCCTCCTGTGGAGCAGAAACACCGCAGTATTGTCCAAATTATTTATGATGAGAATCG 979

Qy          235 gLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuPr 255
                |||||||||||||   ||||||:::|||   |||||||||||||||:::|||||||||||
Db          980 GAAAAAAGCAGAAGAAGCTCATAAAATTTTTGAAGGTCTTGGCCCAAAAGTTGAACTGCC 1039

Qy          255 oLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAl 275
                |||||||||||||||||||||||||:::   ||||||||||||||||||   ||||||  
Db         1040 ACTGTATAACCAGCCATCAGATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGT 1099

Qy          275 aMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLy 295
                 |||||||||||||||||||||:::||||||||||||||||||||||||||||||   ::
Db         1100 GATGAGGAAAAAACTCATTTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGA 1159

Qy          295 sGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluAr 315
                :||||||   ||||||||||||||||||||||||||||||   ||||||||||||:::||
Db         1160 ACAAAAAATCTGCCAGCGTTATGATCAGCTCATGGAGGCATGGGAGAAAAAAGTGGACAG 1219

Qy          315 gIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLy 335
                ||||||||||||||||||||||:::|||||||||||||||   |||||||||||||||||
Db         1220 AATAGAAAATAATCCTCGGAGGAAAGCTAAAGAAAGCAAAACAAGGGAATACTATGAAAA 1279

Qy          335 sGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGl 355
                |||||||||||||||||||||||||||||||   |||||||||   |||   ||||||||
Db         1280 GCAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGG 1336

Qy          355 yGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIl 375
                ||||||||||:::|||||||||   :::   ||||||||||||||||||:::||||||||
Db         1337 GCAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAAT 1396

Qy          375 eIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIl 395
                ||||||||||||||||||||||||||||   |||||||||||||||||||||:::|||||
Db         1397 TATTGATGGGCTCTCTGAGCAGGAGAATAATGAGAAACAAATGCGGCAGCTCTCTGTGAT 1456

Qy          395 eProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLe 415
                ||||||||||::::::||||||:::|||:::|||:::|||||||||||||||||||||||
Db         1457 TCCACCTATGATGTTTGATGCAGAACAAAGACGAGTCAAGTTCATTAACATGAATGGGCT 1516

Qy          415 uMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGl 435
                ||||   ||||||||||||||||||||||||||||||   ||||||:::|||::::::  
Db         1517 TATGGAGGACCCTATGAAAGTGTATAAAGATAGGCAGTTTATGAATGTTTGGACTGACCA 1576

Qy          435 nGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAl 455
                 |||||||||   |||::::::||||||:::|||||||||||||||||||||||||||||
Db         1577 TGAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGC 1636

Qy          455 aSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLy 475
                ||||:::||||||||||||:::|||   :::|||||||||||||||||||||||||||||
Db         1637 ATCATACTTGGAGAGGAAGAGTGTTCCTGATTGTGTTTTGTATTACTATTTAACCAAGAA 1696

Qy          475 sAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSerGl 494
                ||||||||||||||||:::||||||||||||:::|||   :::|||||||||::::::||
Db         1697 AAATGAGAATTATAAAGCCCTCGTCAGAAGGAATTATGGGAAACGCAGAGGCAGAAACCA 1756

Qy          494 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProAr 514
                ||||                                                :::   ||
Db         1757 GCAA------------------------------------------------ATTGCTCG 1768

Qy          514 gSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGl 534
```

```
                |   |||||||||||||||   |||:::::::::|||   :::|||||||||   |||::
Db         1769 ACCCTCGCAAGAAGAAAAAGTAGAAGAAAAAGAAGAGGATAAAGCAGAAAAAAACAGAAAA 1828

Qy          534 uLys 535
                :|||
Db         1829 AAAA 1832

RESULT 8
BC029627
LOCUS       BC029627                1835 bp    mRNA    linear   HTC 07-AUG-2002
DEFINITION  Mus musculus, Similar to nuclear receptor co-repressor 1, clone
            IMAGE:4501757, mRNA.
ACCESSION   BC029627
VERSION     BC029627.1  GI:20988194
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1835)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAY-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
            Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
            Richards, S., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 31 Row: k Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1835
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4501757"
                     /tissue_type="Eye, retina, mouse strain C57Bl\6"
                     /clone_lib="NIH_MGC_94"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:              5.37e-85        Length:         1835
Score:                  1626.50         Matches:        342
Percent Similarity:     73.43%          Conservative:   56
Best Local Similarity:  63.10%          Mismatches:     93
Query Match:            12.31%          Indels:         51
DB:                     3               Gaps:           12

US-09-522-753-5 (1-2517) x BC029627 (1-1835)

Qy           16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
                :::   |||||||||   ||||||:::   |||               |||         
Db          282 CAAAGTCGTTATCCTTCACATTCTGTCCAGTATACCTTTCCCAGTACCCGACACCAGCAG 341
```

```
Qy        36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
             :::      :::   :::|||:::      |||               ||||||   |||
Db       342 GAATTTGCAGTTCCTGACTACCGCTCTTCTCATATTGAAGTTAGCCAGGCATCACAGCTT 401

Qy        54 SerProGlySerIleIleGlnProGln---ArgArgArgProSerLeuLeuSerGluPhe 72
                               |||   |||   ||||||||||||||||||||||||||||||
Db       402 TTACAGCAGCAGCAGCAGCAGCAGCAGCTTCGAAGACGACCTTCCTTGCTTTCAGAATTT 461

Qy        73 GlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSer---HisSerTyr 91
                ||||||::::::|||   ||||||               |||      ||||||
Db       462 CACCCGGGTTCTGACAGGCCTCAAGAAAGGAGAACTGGATATGAACAGTTTCACTCA--- 518

Qy        92 LeuProGluLeuGlyLysSerGluMetGluPhe------IleGluSerLysArgProArg 109
                         |||   |||   ::::::         :::||||||||||||||||||
Db       519 ------------GGACCCTCACCGGTGGATCATGATTCCTTGGAGTCCAAGCGCCCACGC 566

Qy       110 LeuGluLeuLeuProAsp---------------------------------ProLeuLeu 118
             ||||||   :::   |||                                 ||||||:::
Db       567 CTGGAGCCAGTTTCTGATGCCCATTTCCAGCGTGTTAGTGCTGCGGTTTTACCTTTAGTT 626

Qy       119 ArgProSerProLeuLeuAlaThrGlyGlnProAlaGly---------SerGluAspLeu 135
                                           |||   |||         |||   |||
Db       627 CAC---------------------TCGCTGCCAGAAGGCTTGAGGTCGTCTGCAGATGCT 665

Qy       136 ThrLysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProPro 154
                ||||||   :::         |||   |||   |||   ||||||   :::
Db       666 AAGAAGGATTCAGCATTTGGAAGCAAACATGAAGCTCCATCCTCTCCTTTGGCTGGGCAA 725

Qy       155 HisThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGln 174
                         :::            |||   :::||||||||||||||||||||||||
Db       726 CCATGTGGAGATGACCAAAATGCTTCACCTTCAAAGCTTTCAAAGGAGGAGTTAATACAG 785

Qy       175 AsnMetAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLys 194
             :::||||||||||||||||||||||||      |||||||||||||||   |||||||||
Db       786 AGTATGGACCGGGTAGACCGAGAGATTGCAAAAGTAGAGCAGCAGATCCTTAAATTGAAA 845

Qy       195 LysLysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProVal 214
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       846 AAGAAACAACAACAGCTAGAAGAAGAAGCAGCTAAGCCCCCTGAACCAGAGAAGCCTGTG 905

Qy       215 SerProProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsn 234
             ||||||||||||:::|||   ||||||||||||:::||||||||||||||||||||||||
Db       906 TCCCCTCCTCCTGTGGAACAAAAACACCGTAGTATTGTCCAAATTATTTATGATGAGAAT 965

Qy       235 ArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeu 254
             |||||||||||||||   ||||||:::|||   |||||||||||||||:::|||||||||
Db       966 CGGAAAAAAGCAGAAGAAGCTCATAAAATATTTGAAGGTCTTGGCCCAAAAGTTGAACTG 1025

Qy       255 ProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGln 274
             |||||||||||||||||||||||||||:::   ||||||||||||||||||   ||||||
Db      1026 CCACTCTATAACCAGCCATCAGATACCAAGGTGTACCATGAGAACATCAAGACAAACCAG 1085

Qy       275 AlaMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrp 294
                |||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db      1086 GTGATGAGGAAAAAACTCATTTTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGG 1145

Qy       295 LysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGlu 314
             :::||||||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db      1146 GAACAAAAAATCTGCCAGCGCTATGATCAGCTCATGGAAGCGTGGGAGAAAAAAGTGGAC 1205

Qy       315 ArgIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGlu 334
             ||||||||||||||||||||||||:::|||||||||||||||   |||||||||||||||
Db      1206 AGAATAGAAAATAATCCCCGGAGGAAAGCAAAAGAAAGCAAAACAAGGGAATACTATGAG 1265

Qy       335 LysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgVal 354
             |||||||||||||||||||||||||||||||||   |||||||||   |||   ||||||
Db      1266 AAGCAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTT 1322

Qy       355 GlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGlu 374
             ||||||||||||:::|||||||||   :::   ||||||||||||||||||:::||||||
Db      1323 GGTCAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAA 1382

Qy       375 IleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaVal 394
             ||||||||||||||||||||||||||||||   |||||||||||||||||||||:::|||
Db      1383 ATTATTGATGGTCTTTCTGAACAGGAGAATAATGAGAAGCAAATGCGTCAGCTTTCTGTG 1442

Qy       395 IleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGly 414
             ||||||||||||::::::||||||:::|||:::|||:::|||||||||||||||||||||
Db      1443 ATTCCACCTATGATGTTTGATGCAGAACAAAGAAGGGTCAAATTCATCAATATGAATGGG 1502

Qy       415 LeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGlu 434
             ||||||   ||||||||||||||||||||||||||||||   ||||||:::|||::::::
Db      1503 CTGATGGAGGATCCAATGAAGGTTTATAAAGACAGACAGTTTATGAATGTTTGGACTGAC 1562

Qy       435 GlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIle 454
                |||||||||   |||::::::||||||:::|||||||||||||||||||||||||||
Db      1563 CATGAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATT 1622

Qy       455 AlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLys 474
             ||||||:::||||||||||||:::|||   :::|||||||||||||||||||||||||||
Db      1623 GCATCCTATTTGGAAAGGAAGAGTGTTCCTGATTGTGTTTTATATTACTATTTAACCAAG 1682

Qy       475 LysAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSer 493
             ||||||||||||||||||:::||||||||||||:::|||   :::|||||||||::::::
Db      1683 AAAAATGAGAATTATAAGGCCCTCGTGAGAAGGAATTATGGAAAACGCAGAGGCAGAAAT 1742

Qy       494 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetPro 513
             ||||||                                                :::
Db      1743 CAGCAG------------------------------------------------ATTGCC 1754

Qy       514 ArgSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGlu 533
             |||   |||||||||||||||   |||:::::::::|||   :::|||||||||   |||
Db      1755 CGTCCCTCACAAGAAGAAAAAGTAGAAGAAAAGGAAGAGGATAAAGCAGAAAAAAACAGAA 1814

Qy       534 GluLys 535
             :::|||
Db      1815 AAAAAA 1820

RESULT 9
BC062162
LOCUS       BC062162                1819 bp    mRNA    linear   HTC 14-NOV-2003
DEFINITION  Mus musculus cDNA clone IMAGE:30240594, containing frame-shift
            errors.
ACCESSION   BC062162
VERSION     BC062162.1  GI:38328289
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1819)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE  22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1819)
```

```
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe and Dr. Mina
            cDNA Library Preparation: Invitrogen Corp
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 131 Row: c Column: 2
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1819
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30240594"
                     /tissue_type="Jaw bone, mouse, day 10.5 to 11.5 (CD-1),
                     developing maxilla and mandibula tissue containing
                     undifferentiated progenitor cells for muscle, dermis,
                     epidermis, skin, membraneous bone, cartilage and teeth"
                     /clone_lib="NIH_MGC_164"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6.1"
ORIGIN

Alignment Scores:
Pred. No.:              6.94e-85      Length:        1819
Score:                  1624.50       Matches:       341
Percent Similarity:     73.94%        Conservative:  59
Best Local Similarity:  63.03%        Mismatches:    91
Query Match:            12.29%        Indels:        50
DB:                     3             Gaps:          11

US-09-522-753-5 (1-2517) x BC062162 (1-1819)

Qy        16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
             :::   |||||||||   ||||||:::   |||               |||
Db       282 CAAAGTCGTTATCCTTCACATTCTGTCCAGTATACCTTTCCCAGTACCCGACACCAGCAG 341

Qy        36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
             :::      :::   :::|||:::      |||               ||||||   |||
Db       342 GAATTTGCAGTTCCTGACTACCGCTCTTCTCATATTGAAGTTAGCCAGGCATCACAGCTT 401

Qy        54 SerProGlySerIleIleGlnProGln---ArgArgArgProSerLeuLeuSerGluPhe 72
                               |||   |||   ||||||||||||||||||||||||||||||
Db       402 TTACAGCAGCAGCAGCAGCAGCAGCAGCTTCGAAGACGACCTTCCTTGCTTTCAGAATTT 461

Qy        73 GlnProGlyAsnGluArgSerGlnGluLeu-HisLeuArgProGluSerHisSerTyrLe 92
                ||||||::::::|||   ||||||       |||      :::   ||||||
Db       462 CACCCGGGTTCTGACAGGCCTCAAGAAAGGAGAACTGGATGAACAGTTTCACTCA----- 516

Qy        92 uProGluLeuGlyLysSerGluMetGluPhe------IleGluSerLysArgProArgLe 110
                       |||   |||   ::::::         :::||||||||||||||||||||
Db       517 ----------GGACCCTCACCGGTGGATCATGATTCCTTGGAGTCCAAGCGCCCACGCCT 566

Qy       110 uGluLeuLeuProAsp---------------------------------ProLeuLeuAr 119
             ||||   :::   |||                                 ||||||:::
Db       567 GGAGCCAGTTTCTGATGCCCATTTCCAGCGTGTTAGTGCTGCGGTTTTACCTTTAGTTCA 626

Qy       119 gProSerProLeuLeuAlaThrGlyGlnProAlaGly---------SerGluAspLeuTh 136
                                         |||   |||         |||   |||
Db       627 C---------------------TCGCTGCCAGAAGGCTTGAGGTCGTCTGCAGATGCTAA 665

Qy       136 rLysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProProHi 155
              ||||||   :::         |||   |||   |||   ||||||   :::
Db       666 GAAGGATTCAGCATTTGGAAGCAAACATGAAGCTCCATCCTCTCCTTTGGCTGGGCAACC 725

Qy       155 sThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGlnAs 175
                       :::            |||   :::||||||||||||||||||||||||::
Db       726 ATGTGGAGATGACCAAAATGCTTCACCTTCAAAGCTTTCAAAGGAGGAGTTAATACAGAG 785

Qy       175 nMetAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLy 195
             :||||||||||||||||||||||||      |||||||||||||||   |||||||||||
Db       786 TATGGACCGGGTAGACCGAGAGATTGCAAAAGTAGAGCAGCAGATCCTTAAATTGAAAAA 845

Qy       195 sLysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProValSe 215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       846 GAAACAACAACAGCTAGAAGAAGAAGCAGCTAAGCCCCCTGAACCAGAGAAGCCTGTGTC 905
```

```
Qy       215 rProProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnAr 235
             ||||||||||:::|||   ||||||||||||:::||||||||||||||||||||||||||
Db       906 CCCTCCTCCTGTGGAACAAAAACACCGTAGTATTGTCCAAATTATTTATGATGAGAATCG 965

Qy       235 gLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuPr 255
             |||||||||||||   ||||||:::|||   |||||||||||||||:::|||||||||||
Db       966 GAAAAAAGCAGAAGAAGCTCATAAAATATTTGAAGGTCTTGGCCCAAAAGTTGAACTGCC 1025

Qy       255 oLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAl 275
             |||||||||||||||||||||||||:::   ||||||||||||||||||   ||||||
Db      1026 ACTCTATAACCAGCCATCAGATACCAAGGTGTACCATGAGAACATCAAGACAAACCAGGT 1085

Qy       275 aMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLy 295
              |||||||||||||||||||||:::||||||||||||||||||||||||||||||   ::
Db      1086 GATGAGGAAAAAACTCATTTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGGGA 1145

Qy       295 sGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluAr 315
             :||||||   ||||||||||||||||||||||||||||||   ||||||||||||:::||
Db      1146 ACAAAAAATCTGCCAGCGCTATGATCAGCTCATGGAAGCGTGGGAGAAAAAAGTGGACAG 1205

Qy       315 gIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLy 335
             ||||||||||||||||||||||:::|||||||||||||||   |||||||||||||||||
Db      1206 AATAGAAAATAATCCCCGGAGGAAAGCAAAAGAAAGCAAAACAAGGGAATACTATGAGAA 1265

Qy       335 sGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGl 355
             |||||||||||||||||||||||||||||||   |||||||||   |||   ||||||||
Db      1266 GCAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTTGG 1322

Qy       355 yGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIl 375
             ||||||||||:::|||||||||   :::   ||||||||||||||||||:::||||||||
Db      1323 TCAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAAAT 1382

Qy       375 eIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIl 395
             ||||||||||||||||||||||||||||   |||||||||||||||||||||:::|||||
Db      1383 TATTGATGGTCTTTCTGAACAGGAGAATAATGAGAAGCAAATGCGTCAGCTTTCTGTGAT 1442

Qy       395 eProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLe 415
             ||||||||||::::::||||||:::|||:::|||:::|||||||||||||||||||||||
Db      1443 TCCACCTATGATGTTTGATGCAGAACAAAGAAGGGTCAAATTCATCAATATGAATGGGCT 1502

Qy       415 uMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGl 435
             ||||   ||||||||||||||||||||||||||||||   ||||||:::|||::::::
Db      1503 GATGGAGGATCCAATGAAGGTTTATAAAGACAGACAGTTTATGAATGTTTGGACTGACCA 1562

Qy       435 nGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAl 455
              |||||||||   |||::::::||||||:::|||||||||||||||||||||||||||||
Db       1563 TGAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATTGC 1622

Qy        455 aSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLy 475
              ||||:::||||||||||||:::|||   :::|||||||||||||||||||||||||||||
Db       1623 ATCCTATTTGGAAAGGAAGAGTGTTCCTGATTGTGTTTTATATTACTATTTAACCAAGAA 1682

Qy        475 sAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSerGl 494
              ||||||||||||||||:::||||||||||||:::|||   :::|||||||||::::::||
Db       1683 AAATGAGAATTATAAGGCCCTCGTGAGAAGGAATTATGGAAAACGCAGAGGCAGAAATCA 1742

Qy        494 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProAr 514
              |||||||                                             :::   ||
Db       1743 GCAGCAG---------------------------------------------ATTGCCCG 1757

Qy        514 gSerSerGlnGluGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGl 534
              |   |||||||||||||||   |||:::::::::|||   :::|||||||||::::::::
Db       1758 TCCCTCACAAGAAGAAAAAGTAGAAGAAAAGGAAGAGGATAAAGCAGAAAAAAAAAAAAA 1817

Qy        534 u 534
              :
Db       1818 A 1818

RESULT 10
BM909096
LOCUS       BM909096                1093 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6611948 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5431719
            5', mRNA sequence.
ACCESSION   BM909096
VERSION     BM909096.1  GI:19359475
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1093)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1904  row: g  column: 16
            High quality sequence stop: 741.
FEATURES             Location/Qualifiers
     source          1. .1093
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5431719"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              3.58e-85      Length:         1093
Score:                  1624.00       Matches:        331
Percent Similarity:     92.78%        Conservative:   3
Best Local Similarity:  91.94%        Mismatches:     14
```

```
Query Match:              12.29%         Indels:          12
DB:                       5              Gaps:            3

US-09-522-753-5 (1-2517) x BM909096 (1-1093)

Qy       2151 HisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSer 2170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 CACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCTCTACTCCTTCCCTGGGGCCAGC 78

Qy       2171 CysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAspHis 2190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 TGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCCCCCGGACCAT 138

Qy       2191 GlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsn 2210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        139 GGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCCAGAGCCAAAC 198

Qy       2211 LysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProProGluGly 2230
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        199 AAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCCACCGGAGGGC 258

Qy       2231 MetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGlu 2250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        259 ATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCGGGATGGGGAA 318

Qy       2251 GlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAla 2270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        319 CAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCAGCCGCCAGCC 378

Qy       2271 PhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIle 2290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 TTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAAGCAAGAGATC 438

Qy       2291 AsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly 2310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 AACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAGCCAGCCTGGG 498

Qy       2311 ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGln 2330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 ACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTATAGAAGCCAG 558

Qy       2331 AlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeu 2350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAGAAAGGCACTC 618

Qy       2351 MetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnPro 2370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 ATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGCTTTTAACCCT 678

Qy       2371 LeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSer 2390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 CTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGACGGACGGAGT 738

Qy       2391 AspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSer 2410
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        739 GACCACACACTCACCTCGCCAGGTGGCGGCGGGAAAGCCAAGGTCTCTGGCAGACCCAGC 798

Qy       2411 SerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerVal 2430
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        799 AGCCGAAAAGCCAAGTCCCCGGNCCCCGGGCTGGCATCTGGGGACCGGCCACCCTCTGTC 858

Qy       2431 SerSerValHis-SerGluGlyAspCys-AsnArgArgThrProLeu-ThrAsnArgVal 2449
              |||||||||||| ||||||||||||    |||||||||||||||||| ||||||||||||
Db        859 TCCTCAGTGCACTTCGGAGGGAGACTGGAAACCGCCGGACGCCGCTCAACAAACCGCGTG 918

Qy       2450 -TrpGlu-AspArgProSerSerAlaGly-SerThrPro-PheProTyrAsnProLeuIl 2468
               |||||| |||||||||   ||||||||| |||   ||| |||||||||   |||   ||
Db        919 GTGGGAAGGACAGGCCCTTCTCCGCAGGTTTCCCCGCCATTTCCCCTACACCCCCTGGAT 978

Qy       2468 eMetArgLeuGlnAlaGlyValMet------AlaSerProProProProGlyLeuProAl 2486
              |||||||:::      |||   :::         |||   ||||||   |||||||||
Db        979 CATGCGGGTTGACAGGGGGGGTGTTCATGGGTTTTCCCAACCCCCACGGGGCCTCCCCCG 1038
```

```
Qy       2486 aGly---SerGlyProLeuAla---GlyProHis-HisAlaTrp 2498
               |||   |||||||||   :::   ||||||||| |||||||||
Db       1039 CGGGGAAAGCGGGCCCCCTTCCCGTGGCCCCCACCCACGCCTGG 1082

RESULT 11
BC026623
LOCUS       BC026623                1996 bp    mRNA    linear   HTC 20-SEP-2002
DEFINITION  Mus musculus, Similar to nuclear receptor co-repressor 1, clone
            IMAGE:4981672, mRNA.
ACCESSION   BC026623
VERSION     BC026623.1  GI:20073007
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1996)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 59 Row: m Column: 5
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..1996
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4981672"
                     /tissue_type="Colon, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

Alignment Scores:
Pred. No.:              1.45e-83        Length:         1996
Score:                  1603.00         Matches:        337
Percent Similarity:     73.40%          Conservative:   52
Best Local Similarity:  63.58%          Mismatches:     91
Query Match:            12.13%          Indels:         50
DB:                     3               Gaps:           12

US-09-522-753-5 (1-2517) x BC026623 (1-1996)

Qy         16 GluProArgTyrProProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
              :::   |||||||||   ||||||:::   |||               |||
Db        490 CAAAGTCGTTATCCTTCACATTCTGTCCAGTATACCTTTCCCAGTACCCGACACCAGCAG 549

Qy         36 AspValGlyLeuLeuGluTyrGln------HisHisSerArgAspTyrAlaSerHisLeu 53
              :::      :::   :::|||:::      |||               ||||||   |||
Db        550 GAATTTGCAGTTCCTGACTACCGCTCTTCTCATATTGAAGTTAGCCAGGCATCACAGCTT 609

Qy         54 SerProGlySerIleIleGlnProGln---ArgArgArgProSerLeuLeuSerGluPhe 72
                                |||   |||   ||||||||||||||||||||||||||||||
Db        610 TTACAGCAGCAGCAGCAGCAGCAGCTTCGAAGACGACCTTCCTTGCTTTCAGAATTT 669

Qy         73 GlnProGlyAsnGluArgSerGlnGluLeuHisLeuArgProGluSer---HisSerTyr 91
                 ||||||::::::|||   ||||||               |||      ||||||
Db        670 CACCCGGGTTCTGACAGGCCTCAAGAAAGGAGAACTGGATATGAACAGTTTCACTCA--- 726

Qy         92 LeuProGluLeuGlyLysSerGluMetGluPhe------IleGluSerLysArgProArg 109
                          |||   ||| . ::::::         :::||||||||||||||||||
Db        727 ------------GGACCCTCACCGGTGGATCATGATTCCTTGGAGTCCAAGCGCCCACGC 774

Qy        110 LeuGluLeuLeuProAsp---------------------------------ProLeuLeu 118
              ||||||   :::   |||                                 ||||||:::
Db        775 CTGGAGCCAGTTTCTGATGCCCATTTCCAGCGTGTTAGTGCTGCGGTTTTACCTTTAGTT 834

Qy        119 ArgProSerProLeuLeuAlaThrGlyGlnProAlaGly---------SerGluAspLeu 135
                                            |||   |||         |||   |||
Db        835 CAC---------------------TCGCTGCCAGAAGGCTTGAGGTCGTCTGCAGATGCT 873

Qy        136 ThrLysAspArgSerLeuThrGlyLysLeuGlu---ProValSerProProSerProPro 154
                 ||||||   :::         |||   |||   |||   ||||||   :::
Db        874 AAGAAGGATTCAGCATTTGGAAGCAAACATGAAGCTCCATCCTCTCCTTTGGCTGGGCAA 933

Qy        155 HisThrAspProGluLeuGluLeuValProProArgLeuSerLysGluGluLeuIleGln 174
                          :::            |||   :::||||||||||||||||||||||||
Db        934 CCATGTGGAGATGACCAAAATGCTTCACCTTCAAAGCTTTCAAAGGAGGAGTTAATACAG 993

Qy        175 AsnMetAspArgValAspArgGluIleThrMetValGluGlnGlnIleSerLysLeuLys 194
              :::||||||||||||||||||||||||      |||||||||||||||   |||||||||
Db        994 AGTATGGACCGGGTAGACCGAGAGATTGCAAAAGTAGAGCAGCAGATCCTTAAATTGAAA 1053

Qy        195 LysLysGlnGlnGlnLeuGluGluGluAlaAlaLysProProGluProGluLysProVal 214
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1054 AAGAAACAACAACAGCTAGAAGAAGAAGCAGCTAAGCCCCCTGAACCAGAGAAGCCTGTG 1113

Qy        215 SerProProProIleGluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsn 234
              ||||||||||||:::|||   ||||||||||||:::||||||||||||||||||||||||
Db       1114 TCCCCTCCTCCTGTGGAACAAAAACACCGTAGTATTGTCCAAATTATTTATGATGAGAAT 1173

Qy        235 ArgLysLysAlaGluAlaAlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeu 254
              |||||||||||||||   ||||||:::|||   |||||||||||||||:::|||||||||
Db       1174 CGGAAAAAAGCAGAAGAAGCTCATAAAATATTTGAAGGTCTTGGCCCAAAAGTTGAACTG 1233

Qy        255 ProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGln 274
              |||||||||||||||||||||||||||:::   ||||||||||||||||||   ||||||
Db       1234 CCACTCTATAACCAGCCATCAGATACCAAGGTGTACCATGAGAACATCAAGACAAACCAG 1293

Qy        275 AlaMetArgLysLysLeuIleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrp 294
                 |||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1294 GTGATGAGGAAAAAACTCATTTTATTTTTAAAAGAAGAAATCATGCAAGAAAACAAAGG 1353

Qy        295 LysGlnLysPheCysGlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGlu 314
              :::||||||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db       1354 GAACAAAAAATCTGCCAGCGCTATGATCAGCTCATGGAAGCGTGGGAGAAAAAAGTGGAC 1413

Qy        315 ArgIleGluAsnAsnProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGlu 334
              ||||||||||||||||||||||||:::|||||||||||||||   |||||||||||||||
Db       1414 AGAATAGAAAATAATCCCCGGAGGAAAGCAAAAGAAAGCAAAACAAGGGAATACTATGAG 1473

Qy        335 LysGlnPheProGluIleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgVal 354
              |||||||||||||||||||||||||||||||||   |||||||||   |||   ||||||
Db       1474 AAGCAGTTTCCAGAAATTCGAAAACAAAGAGAACAGCAAGAAAGATTTCAG---CGAGTT 1530

Qy        355 GlyGlnArgGlySerGlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGlu 374
              ||||||||||||:::|||||||||   :::   ||||||||||||||||||:::||||||
Db       1531 GGTCAGAGGGGAGCTGGTCTTTCAGCCACCATTGCTAGGAGTGAGCATGAGATTTCTGAA 1590

Qy        375 IleIleAspGlyLeuSerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaVal 394
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||:::|||
Db       1591 ATTATTGATGGTCTTTCTGAACAGGAGAATAATGAGAAGCAAATGCGTCAGCTTTCTGTG 1650
```

```
Qy        395 IleProProMetLeuTyrAspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGly 414
              ||||||||||||::::::||||||:::|||:::|||:::|||||||||||||||||||||
Db       1651 ATTCCACCTATGATGTTTGATGCAGAACAAAGAAGGGTCAAATTCATCAATATGAATGGG 1710

Qy        415 LeuMetAlaAspProMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGlu 434
              ||||||   ||||||||||||||||||||||||||||||   ||||||:::|||::::::
Db       1711 CTGATGGAGGATCCAATGAAGGTTTATAAAGACAGACAGTTTATGAATGTTTGGACTGAC 1770

Qy        435 GlnGluLysGluThrPheArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIle 454
                 |||||||||   |||::::::||||||:::|||||||||||||||||||||||||||
Db       1771 CATGAAAAGGAGATCTTTAAGGACAAGTTTATCCAGCATCCAAAAAACTTTGGACTAATT 1830

Qy        455 AlaSerPheLeuGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLys 474
              ||||||:::||||||||||||:::|||   :::|||||||||||||||||||||||||||
Db       1831 GCATCCTATTTGGAAAGGAAGAGTGTTCCTGATTGTGTTTTATATTACTATTTAACCAAG 1890

Qy        475 LysAsnGluAsnTyrLysSerLeuValArgArgSerTyr---ArgArgArgGlyLysSer 493
              ||||||||||||||||||:::||||||||||||:::|||   :::|||||||||::::::
Db       1891 AAAAATGAGAATTATAAGGCCCTCGTGAGAAGGAATTATGGAAAACGCAGAGGCAGAAAT 1950

Qy        494 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetPro 513
              |||||||||                                             :::
Db       1951 CAGCAGCAG---------------------------------------------ATTGCC 1965

Qy        514 ArgSerSerGlnGluGluLysAspGluLys 523
              |||   |||||||||||||||   :::|||
Db       1966 CGTCCCTCACAAGAAGAAAAAAAAAAAAAA 1995

RESULT 12
BM555371
LOCUS       BM555371                 965 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6577786 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5470428
            5', mRNA sequence.
ACCESSION   BM555371
VERSION     BM555371.1  GI:18795772
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 965)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1977  row: d  column: 13
            High quality sequence stop: 670.
FEATURES             Location/Qualifiers
     source          1..965
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5470428"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.95e-81        Length:         965
Score:                  1559.00         Matches:        304
Percent Similarity:     95.00%          Conservative:   0
Best Local Similarity:  95.00%          Mismatches:     11
Query Match:            11.80%          Indels:         5
DB:                     4               Gaps:           3

US-09-522-753-5 (1-2517) x BM555371 (1-965)

Qy       2182 AspLeuTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSerGlu 2201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GACCTCTACCTCCCGCCCCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAA 62

Qy       2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGly 2221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 GGGGGCAAGAGGTCTCCAGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGT 122

Qy       2222 IleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ATTGAACCTGTGTCCCCACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTG 182

Qy       2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 TACCCGCTGCTGTACCGGGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCT 242

Qy       2262 ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAla 2281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 CCAGGCAACACCAGCCAGCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCC 302

Qy       2282 MetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGlu 2301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ATGGTCAAGTCCAAGAAGCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAG 362

Qy       2302 ProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGly 2321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 CCTGAATACAATATCAGCCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACCGGA 422

Qy       2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 ACAGGCCTTATGACCTATAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATGGGG 482

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerPro 2361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 CTGGAGGCCATAATTAGAAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCCCCG 542

Qy       2362 ProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMet 2381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CCGCTCAGCGCCAATGCTTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATG 602

Qy       2382 ProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGly 2401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CCCATAACCGCTGCTGACGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGCGGG 662

Qy       2402 LysAlaLysValSerGlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeu 2421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 AAGGCCAAGGTCTCTGGCAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTG 722

Qy       2422 AlaSerGlyAspArgProProSerValSerSerValHisSerGluGlyAspCysAsnArg 2441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 GCATCTGGGGACCGGCCACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAACCGC 782

Qy       2442 ArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySerThrPro 2461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 CGGACGCCGCTCACCAACCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGGTCCACGCCA 842

Qy       2462 PheProTyrAsnProLeuIleMet-ArgLeuGlnAlaGly---ValMetAlaSerProPr 2480
              |||||||||   ||||||||||||    ||||||||||||            ||||||||
Db        843 TTCCCCTACAAACCCCTGATCATGCNNGCTGCAGGCGGGNTGGTCATGGTTTTCCCCACC 902

Qy       2480 oProProGly---LeuProAlaGly------SerGlyProLeuAlaGlyProHisHis 2496
              ||||||||||      |||||||||            |||   ||||||      |||
Db         903 CCCACCGGGGCCTTCCCCCGCGGGGCAGCGGGGGCCCCCCTCGGCTGGGGCCCCCCAC 960

RESULT 13
BM553310
LOCUS       BM553310               1126 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6572642 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467163
            5', mRNA sequence.
ACCESSION   BM553310
VERSION     BM553310.1  GI:18791930
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1968  row: 1  column: 12
            High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1..1126
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5467163"
                     /tissue_type="amelanotic melanoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_41"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              3.18e-80        Length:          1126
Score:                  1540.00         Matches:         314
Percent Similarity:     82.19%          Conservative:    9
Best Local Similarity:  79.90%          Mismatches:      16
Query Match:            11.65%          Indels:          55
DB:                     4               Gaps:            1

US-09-522-753-5 (1-2517) x BM553310 (1-1126)

Qy        1457 LeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArgSerLeu 1476
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           2 CTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAAGCACGACGTACGCTCCCTC 61

Qy        1477 IleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAla 1496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          62 ATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGCC 121

Qy        1497 ArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSer 1516
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         122 CGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAGC 181

Qy        1517 SerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPro 1536
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         182 AGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCCG 241

Qy        1537 ArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArg 1556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         242 CGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACGA 301

Qy        1557 GlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSer 1576
               |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         302 GGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTCC 361

Qy        1577 SerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerPro 1596
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         362 AGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCCG 421

Qy        1597 HisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuArg 1616
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         422 CACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCGG 481

Qy        1617 GlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSer 1636
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         482 GGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTCC 541

Qy        1637 IleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAla 1656
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         542 ATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGCC 601

Qy        1657 ProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspThr 1676
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         602 CCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACACG 661

Qy        1677 AlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHis 1696
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         662 GCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCAC 721

Qy        1697 HisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArg 1716
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         722 CACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAGGGGCCTCTCGCCCCGC 781

Qy        1717 GluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGln 1736
               |||||||||||||||||||||||||||||||||||||||
Db         782 GAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGA--------------------- 820

Qy        1737 ValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAsp 1756

Db         820 ------------------------------------------------------------ 820

Qy        1757 ArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerPro 1776

Db         820 ------------------------------------------------------------ 820

Qy        1777 LeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArg 1796
                        ||||||||||||   ||||||||||||      ||||||||||||||||||
Db         821 ---------GGAGGTCCGACACCCTTGACAAAACCACC-CCCACGTCCTCGTCCGAACGG 870

Qy        1797 GluArgAspArg-AspArgGluArgAsp-ArgAspArgGluArgGluLysSerIleLeuT 1816
               |||   |||||| |||   :::|||::: |||:::   :::||||||||||||||||||
Db         871 GAGCCAGACCGGGGATCCAAAGCGGAACCGGAATTCGCAGCGGGAAAAGTCCATCCTCC 930

Qy        1816 hrSerThrThrThrVal-GluHisAlaProIleTrp-ArgProGlyThrGluGlnSerSe 1835
                 |||||||||   ||| |||   ||||||:::||| :::|||   ||||||   :::::
Db         931 CGTCCACCACCGACGTGGGAGCCCGCACCCCTCTGGAAAACCTGTAACGGAGCCAAACAA 990

Qy        1835 rGlySerSerGlySerSerGlyGlyGlyGly 1845
               :||||||         :::||||||||||||
Db         991 CGGAAGCATCCGGGAAACTGGGGGGGGGGGT 1021

RESULT 14
BM477568
LOCUS       BM477568               1057 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6484946 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554420
            5', mRNA sequence.
ACCESSION   BM477568
VERSION     BM477568.1  GI:18526610
KEYWORDS    EST.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1057)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12272  row: p  column: 05
            High quality sequence stop: 674.
FEATURES             Location/Qualifiers
     source          1. .1057
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5554420"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_85"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              9.02e-77       Length:        1057
Score:                  1480.50        Matches:       308
Percent Similarity:     87.57%         Conservative:  2
Best Local Similarity:  87.01%         Mismatches:    27
Query Match:            11.20%         Indels:        20
DB:                     4              Gaps:          6

US-09-522-753-5 (1-2517) x BM477568 (1-1057)

Qy        929 AlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThrProThr 948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCGAGGGCGGCGACAAGAACCGGCTGCTGTCCCCAAGGCCCAGCCTCCTCACCCCGACT 60

Qy        949 GlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLysGln 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCGACCCCCGGGCCAATGCCTCACCCCAGAAGCCACTGGACCTGAAGCAGCTGAAGCAG 120

Qy        969 ArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArgGluAsp 988
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGAGCGGCTGCCATCCCCCCCATCCAGGTCACCAAAGTCCATGAGCCCCCCCGGGAGGAC 180

Qy        989 AlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGlnProGlu 1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGCTCCCACCAAGCCAGCTCCCCCAGCCCCACCGCCACCGCAAAACCTGCAGCCGGAG 240

Qy       1009 SerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAlaPro 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCGACGCCCCTCAGCAGCCTGGCAGCAGCCCCCGGGGCAAGAGCAGGAGCCCGGCACCC 300

Qy       1029 ProAlaAspLysGlu------------------------AlaPheAlaAlaGluAlaGln 1040
              |||||||||||||||                        |||||||||||||||||||||
Db        301 CCCGCCGACAAGGAGGCAGAGAAGCCTGTGTTCTTCCCAGCCTTCGCAGCCGAGGCCCAG 360

Qy       1041 LysLeuProGlyAspProProCysTrpThrSerGlyLeuProPheProValProProArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGCTGCCTGGGGACCCCCCTTGCTGGACTTCCGGCCTGCCCTTCCCCGTGCCCCCCCGT 420
```

```
Qy       1061 GluValIleLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaProPro 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGGTGATCAAGGCCTCCCCGCATGCCCCGGACCCCTCAGCCTTCTCCTACGCTCCACCT 480

Qy       1081 GlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgProPro 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGTCACCCACTGCCCCTGGGCCTCCATGACACTGCCCGGCCCGTCCTGCCGCGCCCACCC 540

Qy       1101 ThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeuGluArg 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACCATCTCCAACCCGCCTCCCCTCATCTCCTCTGCCAAGCACCCCAGCGTCCTCGAGAGG 600

Qy       1121 GlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSerGluHis 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAAATAGGTGCCATCTCCCAAGGAATGTCGGTCCAGCTCCACGTCCCGTACTCAGAGCAT 660

Qy       1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCAAGGCCCCGGTGGGCCCTGTCACCATGGGGCTGCCCCTGCCCATGGACCCCAAAAAG 720

Qy       1161 LeuAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyPro 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTGGCACCCTTCAGCGGAGTGAAGCAGGAGCAGCTGTCCCCACGGGGCCAGGCTGGGCCA 780

Qy       1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
                 |||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        781 CGGGAGAG-CTGGGGGTGCCACA-GCCCAGGAAGCGTCCGTGCTGAGAGGGACAGCTCTG 838

Qy       1201 GlySerValProGlyGlySerIleThrLysGlyIleProSer---ThrArgValProSer 1219
              ||||||      |||||||||||||||||||||||||||      |||      |||
Db        839 GGCTCAGNTTCCGGCGGGAGCATCACCAAAGGCATTCCCCAGCACACGGGGTGCCCTCGG 898

Qy       1220 AspSer-AlaIleThrTyrArg-GlySerIleThrHisGly---ThrProAlaAspValL 1238
              ||||||    |||||||||||| ||||||||||||||||||            |||||||
Db        899 GACAGCGNCCATCACATACCGCGGGTTCCATCACCCACGGNNACGCCCAGCTGGACGTCC 958

Qy       1238 euTyrLysGlyThr---IleThrArgIle---IleGlyGluAspSerProSerArgLeuA 1256
              ||   :::|||      |||||||||            :::|||   |||      |||
Db        959 TGGA-CAAGGCNCACCCATCACCAGGNATCATCGGGCGAAGGACAGNCCGGAGTCGCTTG 1017

Qy       1256 spArg---------GlyArgGluAspSerLeuProLys 1265
                            ||||||            ||||||
Db       1018 GAAACNCCCGNCCGGGAAGGAACAGCCCTGGCCCCAAA 1055

RESULT 15
BM423558
LOCUS       BM423558                1051 bp    mRNA    linear   EST 29-JAN-2002
DEFINITION  AGENCOURT_6399411 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5516745
            5', mRNA sequence.
ACCESSION   BM423558
VERSION     BM423558.1  GI:18391770
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1051)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2017  row: n  column: 10
            High quality sequence stop: 667.
FEATURES             Location/Qualifiers
     source          1. .1051
```

```
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:5516745"
                    /tissue_type="amelanotic melanoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /clone_lib="NIH_MGC_41"
                    /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                    into EcoRI/XhoI sites using the following 5' adaptor:
                    GGCACGAG(G). Library constructed by Ling Hong in the
                    laboratory of Gerald M. Rubin (University of California,
                    Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                    Superscript II RT (Life Technologies). Note: this is a
                    NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:               1.02e-76      Length:         1051
Score:                   1479.50       Matches:        305
Percent Similarity:      78.95%        Conservative:   10
Best Local Similarity:   76.44%        Mismatches:     24
Query Match:             11.20%        Indels:         60
DB:                      4             Gaps:           4

US-09-522-753-5 (1-2517) x BM423558 (1-1051)

Qy       1457 LeuLysTyrAspThrGlyAlaSerThrThr-GlySerLysLysHisAspValArgSerLe 1476
              ||||||||||||:::||||||||||||||| |||||||||||||||||||||||||||||
Db          2 CTCAAGTACGACTCCGGCGCGTCCACCACTGGGCTCCAAAAAGCACGACGTACGCTCCCT 61

Qy       1476 uIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAlaAspAl 1496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCTGGATGTGATGGCCGACGC 121

Qy       1496 aArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSe 1516
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAGCCGGCCAGGGACCGCCAG 181

Qy       1516 rSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysPr 1536
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGTGCCTGAGCTGGGCAAGCC 241

Qy       1536 oArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProAr 1556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GCGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTTTGCCGGCCACCTCCCACG 301

Qy       1556 gGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeuSerSe 1576
              ||||||||||||||||   |||||||||||||||||||||||||||||||||||||||||
Db        302 AGGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCAGGAGGGCAGCCTTTCGTC 361

Qy       1576 rSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerPr 1596
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCGTGAGATCGCCAAGTCCCC 421

Qy       1596 oHisSerThrValProGluHisHisProHisProIleSerProTyrGluHisLeuLeuAr 1616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCCCTATGAGCACCTGCTTCG 481

Qy       1616 gGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspProThrSe 1636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCTGGCCTTCGACCCCACCTC 541

Qy       1636 rIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisLeuAl 1656
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 CATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTACCTGCCCCGACACCTGGC 601

Qy       1656 aProAsnProThrTyrProHisLeuTyrProProTyrLeuIleArgGlyTyrProAspTh 1676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCATCCGCGGCTACCCCGACAC 661

Qy       1676 rAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHi 1696
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 GGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACATCACCTCGCAGCAGATGCA 721

Qy       1696 sHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProAr 1716
              |||||||   |||||||||||||||||||||||||||||||||:::||||||||||||||
Db        722 CCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCTGAAGGGGCTCTCGCCCCG 781

Qy       1716 gGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGl 1736
              |||||||||||||||||||||||||||||||||||||
Db        782 CGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCC----------------------- 818

Qy       1736 nValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAlaMetAs 1756

Db        818 ------------------------------------------------------------ 818

Qy       1756 pArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSerSerPr 1776

Db        818 ------------------------------------------------------------ 818

Qy       1776 oLeuSerPro-GlyGlyProThrHis-LeuThrLysProThrThrThrSerSerSerGlu 1795
                     ||| |||   ||||||||| ||||||||||||            |||||||||
Db        819 -------CCGAGGAAGGCCAACCCACCTTGACCAAACCACCCCCACGTCCCTCGTCCGAG 871

Qy       1796 ArgGluArgAspArg---AspArgGluArgAspArgAspArgGlu-ArgGluLysSerIl 1814
              |||||||||:::|||      :::            ||||||||| ||||||:::|||||
Db        872 CGGGAGCGAGAACGGGGATCCAAAGGCGGGACCGGGGATCGGGAACCGGGAAAGGTCCAT 931

Qy       1814 eLeuThr---SerThrThrThrValGluHisAlaProIleTrp-ArgProGlyThrGluG 1833
              ||||                     :::||||||||||||||| :::|||||||||:::
Db        932 CCTTCCGTTCCCCCCCAACGGGGGGAAGCACGCCCCCATCTGGAAAACCTGGTACCCAAC 991

Qy       1833 lnSerSerGlyGlySerSerGlySer---------SerGlyGlyGlyGly 1845
                         :::   |||            :::||||||||||||
Db        992 CCAAACAACGGGCACCAGGGCACTTACACATTGCGGGGGGGGGGGG 1038

Search completed: November 12, 2005, 03:26:38
Job time : 24452 secs
```

THIS PAGE BLANK (USPTO)